



TUGAS AKHIR – SS141501

**PEMETAAN TINGKAT KONSUMSI ENERGI
BAHAN BAKAR MINYAK (BBM) KABUPATEN/
KOTA DI JAWA TIMUR DENGAN MODEL-BASED
CLUSTERING**

**MUHAMMAD AZHAR
NRP 1313 100 083**

**Dosen Pembimbing
Raden Mohamad Atok, Ph.D
Dr.rer.pol. Dedy Dwi Prastyo**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**



TUGAS AKHIR – SS141501

**PEMETAAN TINGKAT KONSUMSI ENERGI
BAHAN BAKAR MINYAK (BBM) KABUPATEN/
KOTA DI JAWA TIMUR DENGAN MODEL-BASED
CLUSTERING**

**MUHAMMAD AZHAR
NRP 1313 100 083**

**Dosen Pembimbing
Raden Mohamad Atok, Ph.D
Dr.rer.pol. Dedy Dwi Prastyo**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**



FINAL PROJECT – SS 141501

**MAPPING THE LEVEL OF FUEL OIL ENERGY
CONSUMPTION IN REGENCY/CITY IN EAST
JAVA USING MODEL-BASED CLUSTERING**

MUHAMMAD AZHAR

NRP. 1313 100 083

Supervisor

Raden Mohamad Atok, Ph.D

Dr.rer.pol. Dedy Dwi Prastyo

UNDERGRADUATE PROGRAMME

DEPARTMENT OF STATISTICS

FACULTY OF MATHEMATICS AND NATURAL SCIENCES

INSTITUT TEKNOLOGI SEPULUH NOPEMBER

SURABAYA 2017

LEMBAR PENGESAHAN

PEMETAAN TINGKAT KONSUMSI ENERGI BAHAN BAKAR MINYAK (BBM) KABUPATEN/KOTA DI JAWA TIMUR DENGAN MODEL-BASED CLUSTERING

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada
Program Studi Sarjana Departemen Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember

Oleh :
Muhammad Azhar
NRP. 1313 100 083

Disetujui oleh Pembimbing Tugas Akhir:

Raden Mohamad Atok, Ph.D

NIP: 19710915 199702 1 001

Dr.rer.pol. Dedy Dwi Prastyo.

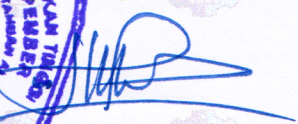
NIP: 19831204 200812 1 002

()

()



Mengetahui,
Kepala Departemen


Dr. Suhartono

NIP. 19710929 199512 1 001

SURABAYA, JULI 2017

PEMETAAN TINGKAT KONSUMSI ENERGI BAHAN BAKAR MINYAK (BBM) KABUPATEN/KOTA DI JAWA TIMUR DENGAN MODEL-BASED CLUSTERING

Nama Mahasiswa : Muhammad Azhar
NRP : 1313 100 083
Departemen : Statistika
Dosen Pembimbing 1 : Raden Mohamad Atok, Ph.D
Dosen Pembimbing 2 : Dr.rer.pol. Dedy Dwi Prastyo
Abstrak

Energi merupakan hal yang sangat dibutuhkan dalam kehidupan, hampir semua sektor kehidupan tidak bisa dipisahkan dari sektor energi. Salah satu jenis energi yang paling banyak dibutuhkan oleh masyarakat adalah energi yang berasal dari minyak bumi khususnya Bahan Bakar Minyak (BBM). Provinsi Jawa Timur merupakan Provinsi dengan jumlah konsumsi BBM terbesar kedua di Indonesia. Berdasarkan jumlah konsumsi BBM di Jawa Timur tersebut, terdapat perbedaan tingkat konsumsi BBM pada Kabupaten/Kota sehingga perlu dilakukan pemetaan/pengelompokkan tingkat konsumsi BBM pada Kabupaten/Kota di Jawa Timur. Pengelompokkan dilakukan dengan metode model-based clustering yaitu metode pengelompokkan yang berdasarkan distribusi peluang dari suatu populasi. Penentuan jumlah kelompok optimal yang terbentuk berdasarkan nilai Integrated Completed Likelihood (ICL) terbesar. Kelompok yang terbentuk akan dilakukan uji kesamaan kelompok untuk mengetahui perbedaan karakteristik pada setiap kelompok. Data yang digunakan adalah jumlah konsumsi BBM Kabupaten/Kota di Jawa Timur tahun 2011-2015 yang dibagi menjadi 20 subset data menurut karakteristik pengelompokkan tertentu. Hasil penelitian menunjukkan bahwa pada beberapa subset data, jumlah kelompok optimal yang terbentuk (G) sebanyak 2 kelompok yang memiliki karakteristik yang berbeda satu sama lain.

Kata Kunci : *Integrated Completed Likelihood (ICL), Konsumsi BBM, Model-Based Clustering, Pengelompokkan*

(halaman ini sengaja dikosongkan)

MAPPING THE LEVEL OF FUEL OIL ENERGY CONSUMPTION IN REGENCY/CITY IN EAST JAVA USING MODEL-BASED CLUSTERING

Name of Student : Muhammad Azhar
NRP : 1313 100 083
Department : Statistics
Supervisor 1 : Raden Mohamad Atok, Ph.D
Supervisor 2 : Dr.rer.pol. Dedy Dwi Prastyo

Abstract

Energy is a very important thing in life, almost all sectors of life can not be separated from the energy sector. One of the most common types of energy needed by society is energy derived from petroleum, especially Fuel Oil (BBM). East Java Province is a Province with the second largest fuel oil consumption in Indonesia. Based on the amount of fuel consumption in East Java, there are different levels of fuel consumption in the Regency/City so it is necessary for mapping/clustering the level of fuel oil consumption in the Regency/City in East Java. Clustering is done by model-based clustering method, it is clustering method based on the probability distribution of a population. Determination of the optimal number of groups formed based on the largest Integrated Completed Likelihood (ICL) value. The group formed will be group equality test to know the characteristic differences in each group. The data used is the amount of fuel consumption of Regency/City in East Java in 2011-2015 which is divided into 20 subsets of data according to certain clustering characteristics. The results showed that in some subset of data, the number of optimal groups formed (G) as many as 2 groups that have different characteristics from each other.

Keywords : Clustering, Fuel Oil Consumption, Integrated Completed Likelihood (ICL), Model-Based Clustering

(halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan rahmat, taufik, serta hidayah-Nya sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul **“Pemetaan Tingkat Konsumsi Energi Bahan Bakar Minyak (BBM) Kabupaten/Kota di Jawa Timur dengan Model-Based Clustering”**.

Tugas Akhir dapat terselesaikan dengan baik dan lancar tidak lepas dari banyaknya bantuan yang diberikan oleh berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada :

1. Raden Mohamad Atok, Ph.D dan Dr.rer.pol. Dedy Dwi Prastyo selaku dosen pembimbing yang telah membimbing dan memberikan dukungan yang sangat besar bagi penulis.
2. Dr. Suhartono dan Dr. Agus Suharsono selaku dosen penguji yang memberikan banyak saran kepada penulis.
3. Ibu (Khusnul Isnaini) dan Ayah (Ribut Riyanto) atas doa, nasehat, dan dukungan yang selalu diberikan kepada penulis.
4. PT. Pertamina (Persero) MOR V Surabaya, Badan Pendapatan Daerah Provinsi Jawa Timur, dan Badan Pusat Statistik Provinsi Jawa Timur yang telah memberikan dukungan data kepada penulis.
5. Teman-teman seperjuangan $\Sigma 24$ yang mengajarkan kebersamaan dan rasa kekeluargaan selama masa perkuliahan.
6. Semua pihak yang telah memberikan dukungan yang tidak dapat penulis sebutkan satu persatu.

Pembuatan laporan Tugas Akhir ini masih jauh dari kesempurnaan, besar harapan dari penulis untuk menerima kritik dan saran yang berguna untuk perbaikan di masa mendatang. Semoga laporan Tugas Akhir ini bermanfaat.

Surabaya, Juli 2017

Penulis

(halaman ini sengaja dikosongkan)

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
TITLE PAGE	iii
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan	4
1.4 Manfaat	4
1.5 Batasan Masalah	4
BAB II TINJAUAN PUSTAKA	
2.1 Statistika Deskriptif	7
2.2 Distribusi t Multivariat	8
2.3 Deteksi <i>Outlier</i> Multivariat	8
2.4 Analisis <i>Cluster</i>	9
2.5 <i>Model-Based Clustering</i>	10
2.6 Model <i>Finite Mixture</i>	12
2.7 Algoritma <i>Expectation Maximization</i> (EM)	13
2.8 <i>Integrated Completed Likelihood</i> (ICL)	14
2.9 <i>Skewness</i> dan <i>Kurtosis</i>	14
2.10 Uji Manova Satu Arah	15
2.11 Konsumsi Bahan Bakar Minyak (BBM)	17
BAB III METODOLOGI PENELITIAN	
3.1 Sumber Data	19

3.2	Variabel Penelitian	19
3.3	Langkah Penelitian	24
BAB IV ANALISIS DAN PEMBAHASAN		
4.1	Deteksi <i>Outlier</i> Multivariat.....	27
4.2	Pengelompokkan menggunakan <i>Model-Based Clustering</i> dengan Kriteria <i>Integrated Completed Likelihood</i> (MBC-ICL).....	29
4.3	Uji Kesamaan Kelompok.....	33
4.4	Analisis Karakteristik Tingkat Konsumsi BBM di Jawa Timur	34
4.4.1	Karakteristik Konsumsi BBM di Jawa Timur Berdasarkan Jumlah Sepeda Motor, Perahu Motor, Penduduk, dan PDRB	34
4.4.2	Karakteristik Konsumsi BBM di Jawa Timur Berdasarkan Jumlah <i>Bus</i> , <i>Truck</i> , Kapal, Penduduk, dan PDRB	38
BAB V PENUTUP		
5.1	Kesimpulan.....	43
5.2	Saran.....	44
DAFTAR PUSTAKA		45
LAMPIRAN.....		49
BIODATA PENULIS		

DAFTAR GAMBAR

	Halaman
Gambar 2.1	Representasi Bentuk <i>Cluster</i> dari Struktur Matriks Varians Kovarians pada <i>Model-Based Clustering</i> 12
Gambar 3.1	Diagram Alir Penelitian..... 25
Gambar 4.1	<i>Plot</i> Jarak Mahalanobis terhadap Jarak Robust <i>Subset</i> Data S_1 2011..... 28
Gambar 4.2	<i>Marginal Contour Plot Subset</i> Data S_1 2015 ... 30
Gambar 4.3	<i>Marginal Contour Plot Subset</i> Data S_2 2015 ... 31
Gambar 4.4	<i>Marginal Contour Plot Subset</i> Data S_3 2015 ... 31
Gambar 4.5	<i>Marginal Contour Plot Subset</i> Data S_4 2015 ... 32
Gambar 4.6	Peta Sebaran Pengelompokkan <i>Subset</i> Data (s_1) Tahun 2011..... 35
Gambar 4.7	Peta Sebaran Pengelompokkan <i>Subset</i> Data (s_1) Tahun 2012..... 35
Gambar 4.8	Peta Sebaran Pengelompokkan <i>Subset</i> Data (s_1) Tahun 2013..... 35
Gambar 4.9	Peta Sebaran Pengelompokkan <i>Subset</i> Data (s_1) Tahun 2014..... 36
Gambar 4.10	Peta Sebaran Pengelompokkan <i>Subset</i> Data (s_1) Tahun 2015..... 36
Gambar 4.11	Peta Sebaran Pengelompokkan <i>Subset</i> Data (s_2) Tahun 2012..... 39
Gambar 4.12	Peta Sebaran Pengelompokkan <i>Subset</i> Data (s_2) Tahun 2013..... 39
Gambar 4.13	Peta Sebaran Pengelompokkan <i>Subset</i> Data (s_2) Tahun 2014..... 40
Gambar 4.14	Peta Sebaran Pengelompokkan <i>Subset</i> Data (s_2) Tahun 2015..... 40

(Halaman ini sengaja dikosongkan)

DAFTAR TABEL

	Halaman
Tabel 3.1 Variabel Penelitian	19
Tabel 3.2 Kombinasi Variabel pada Subset Data Penelitian.....	22
Tabel 3.3 Struktur Subset Data Penelitian Setiap Tahun.....	23
Tabel 3.4 Daftar Kabupaten/Kota di Provinsi Jawa Timur	23
Tabel 4.1 Jumlah Kabupaten/Kota Terdeteksi Sebagai <i>Outlier</i> pada Kuantil 90 Persen	29
Tabel 4.2 Jumlah Kelompok Optimal Metode MBC-ICL.....	32
Tabel 4.3 Hasil Uji Kesamaan Kelompok (Uji Manova) dengan Statistik Uji <i>Wald-Type Test</i>	33
Tabel 4.4 Nilai <i>Mean</i> pada Hasil Pengelompokkan <i>Subset</i> Data (s_1) Tahun 2011-2015.....	37
Tabel 4.5 Nilai <i>Mean</i> pada Hasil Pengelompokkan <i>Subset</i> Data (s_2) Tahun 2012-2015.....	41

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Data Jumlah Konsumsi BBM serta Faktor-Faktor yang Mempengaruhi pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2011 (X1 sampai X16)	49
Lampiran 2 Data Jumlah Konsumsi BBM serta Faktor-Faktor yang Mempengaruhi pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2012 (X1 sampai X16)	51
Lampiran 3 Data Jumlah Konsumsi BBM serta Faktor-Faktor yang Mempengaruhi pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2013 (X1 sampai X16)	53
Lampiran 4 Data Jumlah Konsumsi BBM serta Faktor-Faktor yang Mempengaruhi pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2014 (X1 sampai X16)	55
Lampiran 5 Data Jumlah Konsumsi BBM serta Faktor-Faktor yang Mempengaruhi pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2015 (X1 sampai X16)	57
Lampiran 6 <i>Output Deteksi Outlier Subset Data s₁2011 menggunakan Software R</i>	59
Lampiran 7 Kabupaten/Kota yang terdeteksi sebagai <i>outlier</i> pada <i>subset</i> data s ₁ tahun 2011-2015	60
Lampiran 8 Kabupaten/Kota yang terdeteksi sebagai <i>outlier</i> pada <i>subset</i> data s ₂ tahun 2011-2015	62
Lampiran 9 Kabupaten/Kota yang terdeteksi sebagai <i>outlier</i> pada <i>subset</i> data s ₃ tahun 2011-2015	64

Lampiran 10	Kabupaten/Kota yang terdeteksi sebagai <i>outlier</i> pada <i>subset</i> data s_4 tahun 2011-2015	66
Lampiran 11	<i>Marginal Contour Plot Subset Data</i> s_1 2011	68
Lampiran 12	<i>Marginal Contour Plot Subset Data</i> s_1 2012	68
Lampiran 13	<i>Marginal Contour Plot Subset Data</i> s_1 2013	68
Lampiran 14	<i>Marginal Contour Plot Subset Data</i> s_1 2014	69
Lampiran 15	<i>Marginal Contour Plot Subset Data</i> s_1 2015	69
Lampiran 16	<i>Marginal Contour Plot Subset Data</i> s_2 2011	69
Lampiran 17	<i>Marginal Contour Plot Subset Data</i> s_2 2012	70
Lampiran 18	<i>Marginal Contour Plot Subset Data</i> s_2 2013	70
Lampiran 19	<i>Marginal Contour Plot Subset Data</i> s_2 2014	70
Lampiran 20	<i>Marginal Contour Plot Subset Data</i> s_2 2015	71
Lampiran 21	<i>Marginal Contour Plot Subset Data</i> s_3 2011	71
Lampiran 22	<i>Marginal Contour Plot Subset Data</i> s_3 2012	71
Lampiran 23	<i>Marginal Contour Plot Subset Data</i> s_3 2013	72
Lampiran 24	<i>Marginal Contour Plot Subset Data</i> s_3 2014	72

Lampiran 25	<i>Marginal Contour Plot Subset Data</i> s ₃ 2015	72
Lampiran 26	<i>Marginal Contour Plot Subset Data</i> s ₄ 2011	73
Lampiran 27	<i>Marginal Contour Plot Subset Data</i> s ₄ 2012	73
Lampiran 28	<i>Marginal Contour Plot Subset Data</i> s ₄ 2013	73
Lampiran 29	<i>Marginal Contour Plot Subset Data</i> s ₄ 2014	74
Lampiran 30	<i>Marginal Contour Plot Subset Data</i> s ₄ 2015	74
Lampiran 31	<i>Output Nilai ICL pada Subset Data</i> s ₁ 2015	75
Lampiran 32	<i>Output Estimasi Parameter Subset Data</i> s ₁ 2015 pada Jumlah Kelompok (G) = 2 dengan Model CUCC	76
Lampiran 33	<i>Output Nilai ICL pada Subset Data</i> s ₂ 2015	79
Lampiran 34	<i>Output Estimasi Parameter Subset Data</i> s ₂ 2015 pada Jumlah Kelompok (G) = 2 dengan Model CCCC.....	80
Lampiran 35	<i>Output Nilai ICL pada Subset Data</i> s ₃ 2015	82
Lampiran 36	<i>Output Estimasi Parameter Subset Data</i> s ₃ 2015 pada Jumlah Kelompok (G) = 1 dengan Model UUUU.....	83
Lampiran 37	<i>Output Nilai ICL pada Subset Data</i> s ₄ 2015	85
Lampiran 38	<i>Output Estimasi Parameter Subset Data</i> s ₄ 2015 pada Jumlah Kelompok (G) = 1 dengan Model UUUU.....	86

Lampiran 39	Hasil Pengelompokkan Kabupaten/Kota di Jawa Timur Berdasarkan Karakteristik pada <i>Subset</i> Data s_1 tahun 2011-2015	88
Lampiran 40	Hasil Pengelompokkan Kabupaten/Kota di Jawa Timur Berdasarkan Karakteristik pada <i>Subset</i> Data s_2 tahun 2012-2015	90
Lampiran 41	Surat Pernyataan Data.....	92

BAB I

PENDAHULUAN

1.1 Latar Belakang

Energi merupakan hal yang sangat dibutuhkan dalam kehidupan, hampir semua sektor kehidupan (industri, rumah tangga, transportasi, jasa, dan lain-lain) tidak bisa dipisahkan dari sektor energi (Nuryanti, 2007). Salah satu jenis energi yang paling banyak dibutuhkan oleh masyarakat adalah energi yang berasal dari minyak bumi khususnya Bahan Bakar Minyak (BBM). Salah satu faktor yang menyebabkan tingginya kebutuhan terhadap energi Bahan Bakar Minyak (BBM) adalah meningkatnya jumlah transportasi yang digunakan oleh masyarakat. Badan Pusat Statistik mencatat pada tahun 2014, jumlah kendaraan nasional sebesar 114.209.266 unit. Jumlah tersebut meningkat 9,69 persen dari jumlah kendaraan di tahun sebelumnya (BPS, 2016). Adanya pertumbuhan jumlah kendaraan tersebut menyebabkan masyarakat semakin mempunyai ketergantungan terhadap energi Bahan Bakar Minyak (BBM) tersebut. Berdasarkan data Badan Pengatur Hilir Minyak dan Gas Bumi menunjukkan jumlah konsumsi BBM nasional pada tahun 2014 sebesar 46,789 juta kilo liter (BPH Migas, 2017). Berdasarkan jumlah konsumsi BBM nasional pada tahun 2014 tersebut, tercatat pulau jawa merupakan pulau dengan konsumsi BBM paling besar dibandingkan pulau lainnya yaitu sebesar 12,663 juta kilo liter. (Kementerian ESDM, 2014).

Provinsi Jawa Timur pada tahun 2014 merupakan provinsi yang menduduki peringkat kedua konsumsi Bahan Bakar Minyak tertinggi di Indonesia. Jumlah konsumsi BBM di Jawa Timur pada tahun 2014 mencapai 3,129 juta kilo liter (Kementerian ESDM, 2014). Berdasarkan jumlah konsumsi BBM di Jawa Timur tersebut, kota Surabaya merupakan kota dengan jumlah konsumsi BBM tertinggi yaitu sebesar 14 persen dari total konsumsi BBM Provinsi Jawa Timur, selanjutnya pada urutan kedua terdapat Kabupaten Sidoarjo dengan jumlah konsumsi BBM sebesar 8 persen dari total konsumsi BBM Provinsi Jawa Timur. Sedangkan

Kota Mojokerto merupakan kota dengan jumlah konsumsi BBM terendah yaitu hanya sebesar 0,51 persen dari jumlah konsumsi BBM Provinsi Jawa Timur (Pertamina, 2016). Hal tersebut mengindikasikan bahwa terjadi perbedaan tingkat konsumsi BBM antar Kabupaten/Kota di Jawa Timur. Perbedaan tingkat konsumsi BBM tersebut dapat disebabkan oleh kondisi tipologi dan sistem transportasi pada setiap Kota/Kabupaten yang berbeda-beda. Kondisi tipologi kota dapat ditinjau dari beberapa faktor diantaranya: kepadatan penduduk, jumlah penduduk, tataguna lahan, dan PDRB. Sedangkan kondisi sistem transportasi dapat ditinjau berdasarkan panjang jalan, pola jaringan jalan, kondisi jalan, angkutan umum penumpang, angkutan barang, panjang trayek angkutan umum serta kendaraan pribadi (Handajani, 2012). Berdasarkan hal tersebut, diperlukan adanya pemetaan tingkat konsumsi BBM Kabupaten/Kota di Jawa Timur berdasarkan faktor-faktor penyebab terjadinya perbedaan tingkat konsumsi BBM di setiap Kabupaten/Kota tersebut. Dengan adanya pemetaan tingkat konsumsi BBM, PT. Pertamina selaku produsen sekaligus distributor Bahan Bakar Minyak dapat memonitoring serta menyesuaikan tingkat distribusi penjualan BBM dengan tingkat kebutuhan konsumsi BBM pada setiap Kabupaten/Kota di Jawa Timur.

Salah satu metode yang dapat digunakan untuk melakukan pemetaan adalah metode *clustering*. Metode *clustering* adalah metode yang digunakan untuk mengelompokkan objek data ke dalam kelompok/*cluster* yang sama (Kamber, 2006). *Cluster* adalah sekumpulan objek data yang memiliki kesamaan satu sama lain disatukan dalam kelompok yang sama dan tidak memiliki kesamaan dengan objek data yang lain (Hosseini, 2010). Beberapa penelitian menggunakan metode *clustering* untuk tujuan klasifikasi perilaku konsumen dan melakukan pemetaan segmen *customer* diantaranya dilakukan oleh Wahyudi (2010) menggunakan metode *cluster K-Means* untuk mengetahui konsumsi dan perilaku konsumen dalam penggunaan energi Bahan Bakar Minyak untuk kendaraan bermotor di Surakarta. Hasil penelitian tersebut

menunjukkan pengguna kendaraan bermotor di Surakarta dapat dikelompokkan menjadi tiga *cluster* berdasarkan karakteristik usia, jenis kelamin, pendapatan, pekerjaan, jumlah anggota keluarga, dan jumlah konsumsi bahan bakar minyak. Selain itu penelitian yang dilakukan oleh Susanty (2012) menggunakan metode *two stage clustering* untuk mengetahui segmentasi pelanggan produk PT. XYZ. Penelitian tersebut dilakukan melalui dua tahap *clustering* yaitu menggunakan metode *ward* untuk tahap pertama dan menggunakan metode *K-Means* untuk tahap kedua, didapatkan hasil pelanggan produk PT. XYZ memiliki empat segmen yaitu *Platinum Customers*, *Uncertain Lost Customers*, *Uncertain New Customers*, dan *High Frequency Buying Customers*.

Secara umum, beberapa metode *cluster* dapat dikelompokkan berdasarkan algoritma proses yang dilakukan, yaitu teknik yang berdasarkan ukuran jarak sebagai basis pengelompokannya. Metode berbasis ukuran jarak ini terdiri dari metode *cluster* berhierarki dengan penggabungan (*agglomerative*), antara lain metode Ward dan juga metode *cluster* tak berhierarki, misalnya metode *K-means* (Andesberg, 1973). Metode-metode *cluster* tersebut dalam pembentukan kelompok/*clusternya* hanya memperhatikan ukuran jarak antar objek pengamatan tanpa mempertimbangkan aspek statistik lainnya, seperti distribusi data ataupun objek pengamatan pada *cluster* yang saling tumpang tindih. Oleh karena itu diperlukan metode *cluster* yang teknik perhitungannya tidak hanya didasarkan pada ukuran jarak antar objek pengamatannya, namun juga memperhatikan aspek statistik dari data pengamatan. Salah satu alternatif metode *cluster* yang dapat digunakan adalah metode *model-based clustering*. Pardede (2007) melakukan penelitian tentang perbandingan metode *cluster K-Means* dan *model-based*. Berdasarkan penelitian tersebut, didapatkan hasil kesalahan pengelompokan pada metode *model-based clustering* sebesar 3.33 persen (5 pengamatan), sementara kesalahan pengelompokan metode *K-means* sebesar 10.67 persen (16 pengamatan). Hal ini menunjukkan bahwa metode *model-*

based lebih efektif memisahkan *cluster* yang saling tumpang tindih dibandingkan dengan metode *K-means*. Secara umum untuk data pengamatan yang cenderung terdapat *cluster* yang saling tumpang tindih disarankan menggunakan metode *model-based clustering*.

1.2 Rumusan Masalah

Berdasarkan latar belakang, rumusan masalah yang dibuat dalam penelitian ini adalah melakukan pemetaan tingkat konsumsi energi Bahan Bakar Minyak (BBM) Kabupaten/Kota di Jawa Timur tahun 2011-2015 dengan menggunakan metode *model-based clustering*.

1.3 Tujuan

Berdasarkan latar belakang dan rumusan masalah, maka tujuan penelitian ini adalah sebagai berikut :

1. Melakukan pemetaan tingkat konsumsi Bahan Bakar Minyak (BBM) Kabupaten/Kota di Jawa Timur tahun 2011-2015 dengan menggunakan metode *model-based clustering*.
2. Menentukan jumlah kelompok/*cluster* yang terbentuk dari Kabupaten/Kota di Jawa Timur berdasarkan tingkat konsumsi Bahan Bakar Minyak (BBM).
3. Mengetahui perubahan pengelompokkan Kabupaten/Kota di Jawa Timur tahun 2011-2015 berdasarkan tingkat konsumsi Bahan Bakar Minyak (BBM).

1.4 Manfaat

Hasil penelitian ini diharapkan dapat dijadikan sebagai masukan kepada PT. Pertamina (Persero) MOR V Surabaya selaku produsen sekaligus distributor Bahan Bakar Minyak agar dapat memonitoring serta menyesuaikan tingkat distribusi penjualan BBM dengan tingkat kebutuhan konsumsi BBM pada Kabupaten/Kota di Jawa Timur.

1.5 Batasan Masalah

Berdasarkan rumusan masalah yang telah dibuat, maka batasan masalah pada penelitian ini adalah:

1. Data yang digunakan dalam penelitian ini adalah data sekunder tentang jumlah penjualan Bahan Bakar Minyak (BBM) di Kabupaten/Kota di Jawa Timur tahun 2011-2015 yang diambil dari PT. Pertamina (Persero) MOR V Surabaya.
2. Data jumlah penjualan Bahan Bakar Minyak (BBM) yang digunakan dalam penelitian ini adalah data jumlah penjualan BBM yang berasal dari seluruh Stasiun Pengisian Bahan Bakar Umum (SPBU) pada Kabupaten/Kota di Jawa Timur.
3. Metode yang digunakan dalam penelitian ini adalah metode *model-based clustering*.

(Halaman ini sengaja dikosongkan)

BAB II

TINJAUAN PUSTAKA

2.1. Statistika Deskriptif

Statistika deskriptif adalah metode statistika yang berfungsi untuk memberikan gambaran umum tentang penyajian sampel atau populasi. Analisis statistika deskriptif dapat diartikan sebagai metode yang berkaitan dengan mengumpulkan, mengolah, dan menyajikan data sehingga memberikan informasi yang berguna (Walpole, 1995). Data dapat disajikan dalam bentuk grafik atau tabel, untuk ukuran pemusatan data dan ukuran penyebaran data dideskripsikan secara numerik. Ukuran pemusatan data meliputi rata-rata, nilai tengah, dan modus sedangkan ukuran penyebaran data meliputi ragam dan standar deviasi (Walpole, 1995).

Rata-rata (*mean*) adalah nilai yang dapat digunakan untuk memberikan informasi atau gambaran umum dari sekumpulan data. Perhitungan rata-rata dengan menjumlahkan nilai pada data dibagi dengan banyaknya data tersebut. Rumus yang digunakan untuk menghitung rata-rata adalah sebagai berikut (Walpole, 1995).

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \text{ dimana } n = \text{banyak data} \quad (2.1)$$

Nilai maksimal adalah nilai yang paling tinggi atau besar dari sekumpulan data yang telah diurutkan. Nilai minimum adalah nilai yang paling rendah atau kecil dari sekumpulan data yang telah diurutkan. Ragam (*varians*) adalah ukuran yang digunakan untuk melihat seberapa besar penyimpangan data. Perhitungan ragam ditampilkan pada persamaan sebagai berikut (Walpole, 1995).

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \quad (2.2)$$

2.2 Distribusi t Multivariat

Distribusi t multivariat merupakan perluasan dari distribusi t univariat. Distribusi t multivariat mulai banyak digunakan dalam *model-based clustering*, karena banyak distribusi data multivariat yang tidak mengikuti distribusi normal akibat keberadaan *outlier*, sehingga distribusi data menjadi lebih landai. Kemudian dikembangkan distribusi t untuk data multivariat yang diakui memiliki kemampuan mengatasi *outlier* lebih mudah dibandingkan distribusi normal multivariat (Andrews dkk., 2011).

Vektor variabel acak $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_p]^T$ berdistribusi t multivariat dengan derajat bebas ν , vektor rata-rata $\boldsymbol{\mu} = [\mu_1 \ \mu_2 \ \dots \ \mu_p]^T$ dan matriks varians kovarian $\boldsymbol{\Sigma}$ memiliki fungsi kepadatan peluang sebagai berikut :

$$f(\mathbf{x}) = \frac{\Gamma\left(\frac{\nu + p}{2}\right)}{(\pi\nu)^{\frac{p}{2}} \Gamma\left(\frac{\nu}{2}\right) |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \left(1 + \frac{(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})}{\nu}\right)^{-\left(\frac{\nu + p}{2}\right)} ; \nu > 2 \quad (2.3)$$

dengan $\Gamma(\nu) = \int_0^{\infty} x^{\nu-1} e^{-x} dx$ dan ν disebut juga parameter bentuk

(*shape*) karena fungsi pada persamaan (2.5) dapat diturunkan atau dinaikkan dengan memberikan variasi nilai ν tersebut.

Menurut DeCarlo (1997) distribusi t merupakan distribusi yang mempunyai bentuk *leptokurtic* dengan ekor yang lebih panjang (*longer tails*) dan puncak yang lebih tinggi (*higher peak*) daripada distribusi normal. Hal ini berarti ukuran keruncingan (*kurtosis*) distribusi t lebih tinggi dari pada *kurtosis* distribusi normal yaitu lebih dari 3.

2.3 Deteksi *Outlier* Multivariat

Data tidak bersitribusi normal multivariat atau matriks varians- kovariansnya tidak homogen dapat disebabkan ada

observasi yang mempunyai pola berbeda dengan sebagian besar pola data. Observasi tersebut disebut pencilan (*outlier*). Dalam kasus pengelompokan dengan metode *model-based clustering*, *outlier* dapat menjadikan hasil pengelompokan kurang tepat dan penaksir parameter menjadi bias (McLachlan & Peel, 2000). Oleh sebab itu, sangat penting memeriksa keberadaan *outlier*. Metode yang digunakan untuk memeriksa keberadaan *outlier* multivariat adalah perhitungan jarak mahalanobis yang didefinisikan sebagai berikut :

$$MD_i = \left[(\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) \right]^{\frac{1}{2}}, i = 1, 2, \dots, n \quad (2.4)$$

dimana $\bar{\mathbf{x}}$ adalah vektor *mean* sampel dan \mathbf{S} adalah matriks varians kovarians sampel. Suatu observasi dikatakan *outlier* jika nilai jarak mahalanobis lebih besar dari nilai $\sqrt{\chi_{p;1-\alpha/2}^2}$ dimana p adalah derajat bebas (Rousseeuw & Zomeren, 1990).

2.4 Analisis Cluster

Analisis *cluster* merupakan suatu teknik analisis statistik yang ditujukan untuk menempatkan sekumpulan obyek ke dalam dua atau lebih grup berdasarkan kesamaan-kesamaan obyek atas dasar berbagai karakteristik. Analisis *cluster* berusaha menemukan pengelompokan secara alamiah sesuai dengan *sense* peneliti. Penempatan objek dilakukan dengan mencari pola dalam kumpulan data agar hasil yang didapatkan optimal. Ciri sebuah *cluster* yang baik adalah *cluster* yang mempunyai homogenitas (kesamaan) yang tinggi antar anggota dalam satu *cluster* (*within cluster*). Sedangkan antar *cluster* satu dengan yang lainnya (*between cluster*) diharapkan memiliki heterogenitas (perbedaan) yang tinggi (Rencher & Christensen, 2002).

Secara umum dalam analisis *cluster* Metode pengelompokan dibedakan menjadi dua yaitu metode hierarki (*hierarchical clustering*) dan metode non-hierarki (*non-hierarchical clustering*) (Johnson & Wichern, 2007). Dalam analisis *cluster* hierarki, *cluster* dibentuk dengan melakukan pendekatan-pendekatan tanpa menentukan jumlah kelompok terlebih dahulu. Jumlah kelompok

beserta pengelompokannya akan terbentuk dari pendekatan-pendekatan yang dilakukan. Jarak yang digunakan dalam analisis ini adalah jarak *Euclidian* sebagai berikut :

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_p - y_p)^2} \quad (2.5)$$

Beberapa metode penggabungan yang bisa digunakan dalam analisis cluster hirarki antara lain metode pautan tunggal (*single linkage*), pautan lengkap (*complete linkage*), pautan rata-rata (*average linkage*), metode *ward* (*ward's method*), dan metode *centroid* (*centroid method*).

Selanjutnya untuk metode non hirarki dipakai jika banyaknya *Cluster* sudah diketahui dan biasanya metode ini dipakai untuk mengelompokkan data yang berukuran besar (Hair dkk., 2009). Perhatian utama dalam metode non hirarki adalah bagaimana memilih awalan atau inisial *cluster* yang berpengaruh besar terhadap hasil akhir analisis *cluster*. Adapun yang termasuk dalam metode ini adalah metode *K-means*, *K-Modes*, *Density Estimation* dan *Mixtures of Distribution*. Metode *K-means* hanya dapat digunakan dalam data numerik sehingga untuk data yang sifatnya kategorik dikembangkan sebuah metode yang disebut *K-Modes*. *K-modes* memiliki dasar pemikiran yang sama dengan *K-means* namun menggunakan modus data sebagai pusat centroid (Rencher & Christensen, 2002). Sedangkan *Density Estimation* melakukan pengelompokan dengan mencari wilayah yang memiliki kepadatan/*density* yang tinggi. Estimasi *density* dilakukan menggunakan fungsi kernel dan tidak memerlukan asumsi khusus. Kemudian untuk *Mixtures of Distribution* mendasari identifikasi kelompok melalui distribusi probabilitas. Keseluruhan populasi dimodelkan sebagai sebuah *mixture* dari distribusi sehingga dapat dikatakan pengelompokan pada metode ini adalah pengelompokan berdasarkan model dari distribusi probabilitas.

2.5 Model-Based Clustering

Metode pengelompokan objek yang banyak dikenal adalah metode *clustering* hirarki (*single linkage*, *complete linkage*,

average linkage dan *Ward's linkage*) dan *clustering* non hirarki (K-means). Namun metode pengelompokan tersebut tidak mempunyai dasar pengelompokan secara statistik. Adapun metode pengelompokan yang memperhatikan model statistik disebut dengan *model-based clustering* (MBC). Model ini pertama kali digunakan oleh Banfield & Raftery (1993) untuk pengelompokan objek dalam populasi. Asumsi yang digunakan pada *model-based clustering* adalah dalam suatu populasi dapat diambil subpopulasi yang mempunyai distribusi peluang tertentu dan masing-masing subpopulasi mempunyai parameter yang berbeda. Keseluruhan subpopulasi mempunyai distribusi peluang *mixture* dengan proporsi berbeda untuk setiap subpopulasi. Asumsi ini mengarahkan pada model probabilitas matematika yaitu model *finite mixture*. Saat ini penggunaan model *finite mixture* pada *clustering* telah berkembang sangat cepat dan menjadi salah satu metode *clustering* yang populer. Banfield & Raftery (1993) mengembangkan kerangka *model-based clustering* menggunakan dekomposisi *eigenvalue* dari matriks varians kovarians (Σ_g) sebagai berikut :

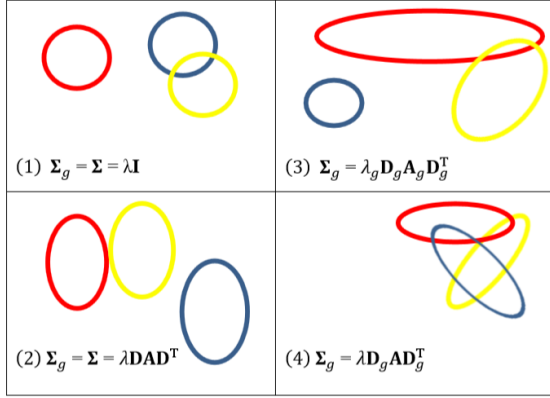
$$\Sigma_g = \lambda_g D_g A_g D_g^T \quad (2.6)$$

dengan :

λ_g adalah nilai skalar yang menunjukkan volume elips.

D_g adalah matriks ortogonal *eigenvector* yang merupakan orientasi dari komponen utama Σ_g .

A_g adalah diagonal matriks dengan elemen-elemen yang proporsional pada *eigenvalue* Σ_g dan menunjukkan kontur dari fungsi kepadatannya. Orientasi, volume dan bentuk dari distribusi dapat diestimasi dari data dan mempunyai bentuk bermacam-macam antar cluster atau dapat saling memotong antar cluster, sebagaimana diilustrasikan pada Gambar 2.1.



Gambar 2.1 Representasi Bentuk *Cluster* dari Struktur Matriks Varians Kovarians pada *Model-Based Clustering*

Keterangan :

1. $\Sigma_g = \Sigma = \lambda \mathbf{I}$, *cluster* yang terbentuk adalah *spherical* (bola) dan mempunyai volume yang sama besar antar *cluster*.
2. $\Sigma_g = \Sigma = \lambda \mathbf{D} \mathbf{A} \mathbf{D}^T$, semua *cluster* yang terbentuk mempunyai bentuk, volume dan orientasi yang sama.
3. $\Sigma_g = \Sigma = \lambda_g \mathbf{D}_g \mathbf{A}_g \mathbf{D}_g^T$, *cluster* yang terbentuk bisa berbeda pada bentuk, volume dan orientasinya.
4. $\Sigma_g = \lambda \mathbf{D}_g \mathbf{A} \mathbf{D}_g^T$ hanya orientasi dari *cluster* terbentuk yang berbeda.

(Agustini, 2017).

2.6 Model *Finite Mixture*

Misalkan vektor variabel acak \mathbf{x} dengan dimensi p berasal dari distribusi *finite mixture* dengan fungsi kepadatan peluang :

$$f(\mathbf{x}|\boldsymbol{\phi}) = \sum_{g=1}^G \pi_g f_g(\mathbf{x}|\boldsymbol{\theta}_g) \quad (2.7)$$

dengan $\boldsymbol{\phi} = (\pi_1, \pi_2, \dots, \pi_G, \theta_1, \theta_2, \dots, \theta_G)$ adalah vektor parameter, $f_g(\mathbf{x}|\boldsymbol{\theta}_g)$ disebut fungsi kepadatan peluang \mathbf{x} dengan parameter kelompok $\boldsymbol{\theta}_g$, G adalah jumlah kelompok dan π_g adalah bobot atau proporsi campuran (*mixing proportion*) dari kelompok ke- g dengan batasan $0 \leq \pi_g \leq 1$, ($g = 1, 2, \dots, G$) dan $\sum_{g=1}^G \pi_g = 1$.

Pada model *finite mixture*, sedikitnya ada dua proses utama yang perlu dilakukan yaitu penaksiran parameter dan pemilihan model terbaik. Penaksiran parameter *model-based clustering* dapat dilakukan dengan algoritma *Expectation Maximization* (EM) (Dempster dkk., 1997), sedangkan pemilihan model terbaik dapat menggunakan kriteria *Integrated Completed Likelihood* (ICL) (Bienarcki dkk., 2000).

2.7 Algoritma *Expectation Maximization* (EM)

Proses iterasi mendapatkan penaksir parameter *model-based clustering* dilakukan dengan algoritma EM untuk mempermudah dalam memperoleh penaksir parameter. Algoritma EM menggunakan iterasi 2 langkah sampai mencapai hasil yang konvergen, yaitu langkah Ekspektasi (*E-step*) dan langkah Maksimisasi (*M-step*).

1. Langkah Ekspektasi (*E-step*)

Langkah ini dilakukan dengan menghitung nilai ekspektasi dari fungsi logaritma likelihood data lengkap yaitu :

$$\ln(L_c(\boldsymbol{\phi}|\mathbf{y}_i)) = \sum_{i=1}^n \sum_{g=1}^G z_{ig} \left[\ln(\pi_g) + \ln(f_g(\mathbf{x}_i|\boldsymbol{\theta}_g)) \right] \quad (2.8)$$

dengan z_{ig} merupakan vektor label pengamatan dalam kelompok ke- g . Berdasarkan ekspektasi dari persamaan (2.10), didapatkan penaksir parameter

$$\hat{z}_{ig} = \frac{\hat{\pi}_g f_g(\mathbf{x}_i|\boldsymbol{\theta}_g)}{\sum_{g=1}^G \hat{\pi}_g f_g(\mathbf{x}_i|\boldsymbol{\theta}_g)} \quad (2.9)$$

2. Langkah Maksimisasi (M-step)

Langkah maksimisasi digunakan untuk memaksimalkan nilai harapan fungsi likelihood data lengkap yaitu :

$$L_c(\boldsymbol{\phi} | \mathbf{y}_i) = \prod_{i=1}^n \prod_{g=1}^G \pi_g^{z_{ig}} \left[f_g(\mathbf{x}_i | \boldsymbol{\theta}_g) \right]^{z_{ig}} \quad (2.10)$$

Persamaan (2.12) tersebut selanjutnya dimaksimisasi nilai harapannya terhadap $\hat{\pi}_g$, $\hat{\boldsymbol{\mu}}_g$ dan $\hat{\boldsymbol{\Sigma}}_g$ dengan nilai \hat{z}_{ig} diperoleh dari langkah Ekspektasi (E-step) (Dempster dkk., 1977).

2.8 Integrated Completed Likelihood (ICL)

Pemilihan *model-based clustering* terbaik bisa menggunakan kriteria *Integrated Completed Likelihood* ICL. Prinsipnya adalah memaksimumkan fungsi likelihood data lengkap $\mathbf{y}_i = (\mathbf{x}_i, \mathbf{z}_i)$ sehingga didapatkan rumus ICL sebagai berikut :

$$ICL_g = \ln(f(\mathbf{y}_i)) - \frac{p}{2} \ln(n) \quad (2.11)$$

dengan $f(\mathbf{y}_i) = f(\mathbf{x}_i, \hat{\mathbf{z}}_i)$ adalah fungsi kepadatan peluang bersama data lengkap, sedangkan p adalah banyaknya parameter, dan n adalah banyaknya observasi (Bienarcki dkk., 2000).

2.9 Skewness dan Kurtosis

Salah satu metode untuk memeriksa distribusi normal multivariat adalah menghitung ukuran *skewness* dan *kurtosis* data. Menurut Mardia (1970) pemeriksaan distribusi normal multivariat menggunakan ukuran multivariate skewness ($\beta_{1,p}$) dan kurtosis data ($\beta_{2,p}$) dengan rumus sebagai berikut :

$$\beta_{1,p} = E \left((\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right)^3 \quad (2.12)$$

$$\beta_{2,p} = E \left((\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right)^2 \quad (2.13)$$

Untuk sampel *random*, ukuran *multivariate skewness* ($b_{1,p}$) dan kurtosis ($b_{2,p}$) adalah

$$b_{1,p} = \frac{1}{n^2} \left[\sum_{i=1}^n \left\{ (\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) \right\}^3 \right] \quad (2.14)$$

$$b_{2,p} = \frac{1}{n} \left[\sum_{i=1}^n \left\{ (\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) \right\}^2 \right] \quad (2.15)$$

Dengan $\mathbf{x}_i = [x_{1i} \ x_{2i} \dots x_{pi}]^T$, $i = 1, 2, \dots, n$, dengan n adalah banyaknya pengamatan $\bar{\mathbf{x}} = [\bar{x}_1 \ \bar{x}_2 \dots \bar{x}_p]^T$ adalah vektor rata-rata sampel dan \mathbf{S} adalah matriks varians kovarians sampel. Data berdistribusi normal jika $\beta_{1,p} = 0$ dan $\beta_{2,p} = p(p+2)$. Hipotesis yang digunakan dalam pengujian distribusi normal dengan nilai *skewness* dan *kurtosis* adalah sebagai berikut.

1. Uji *Skewness*

$$H_0 : \beta_{1,p} = 0$$

$$H_1 : \beta_{1,p} \neq 0$$

$$\text{Dengan statistik uji : } \frac{n}{6} b_{1,p}^2 \sim \chi_{p(p+1)(p+2)/6}^2$$

2. Uji *Kurtosis*

$$H_0 : \beta_{2,p} = p(p+2)$$

$$H_1 : \beta_{2,p} \neq p(p+2)$$

$$\text{Dengan statistik uji : } \frac{b_{2,p} - \frac{n-1}{n+1} p(p+2)}{\sqrt{\frac{8}{n} p(p+2)}} \sim N(0,1)$$

2.10 Uji Manova Satu Arah

Menurut Johnson & Wichern (2007), dalam suatu analisis seringkali terdapat lebih dari satu populasi yang ingin dibandingkan. Sampel acak dikumpulkan dari populasi yang disusun sebagai berikut :

Populasi 1 : $x_{11} \ x_{12} \dots x_{1n_1}$

Populasi 2 : $x_{21} \ x_{22} \dots x_{2n_2}$

⋮

Populasi G : $x_{G1} \ x_{G2} \dots x_{Gn_2}$

Multivariate analysis of variance (manova) digunakan untuk mengetahui apakah vektor rata-rata populasi sama, dan jika tidak sama, komponen rata-rata mana yang berbeda secara nyata. Model Manova untuk membandingkan G vektor rata-rata populasi :

$$x_{gi} = \mu + \tau_g + \varepsilon_{gi}$$

dengan $i = 1, 2, \dots, n_g$ dan $g = 1, 2, \dots, G$, μ vektor rata-rata keseluruhan dan τ_g menyatakan pengaruh perlakuan ke- g . Hipotesis yang digunakan dalam pengujian manova adalah sebagai berikut.

H_0 : $\tau_1 = \tau_2 = \dots = \tau_G = 0$ (tidak ada perbedaan vektor *mean* antar kelompok)

H_1 : Minimal ada satu pasang $\tau_g \neq \tau_{g'}$ dengan g' adalah kelompok selain kelompok g (ada perbedaan vektor *mean* antar kelompok)

Statistik Uji menggunakan *wald-type test* :

$$Q_N(T) = N\bar{X}^T T(\widehat{TV}_N T)^+ T\bar{X} \quad (2.16)$$

dengan $\bar{X} = (\bar{X}_1, \dots, \bar{X}_d)^T$, $\bar{X}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} X_{i,k}$, $\widehat{V}_N = \text{diag} \left(\frac{N}{n_i} \hat{\Sigma}_i \right)$,

$\hat{\Sigma}_i = \frac{1}{n_i - 1} \sum_{k=1}^{n_i} (X_{ik} - \bar{X}_i)(X_{ik} - \bar{X}_i)^T$, $1 \leq i \leq d$, $(T\widehat{V}_N T)^+$ adalah

matriks *Moore-Penrose generalized inverse*, dan $N = \sum_{g=1}^G n_g$

(Bathke dkk., 2016). Keputusan tolak H_0 jika p_value lebih kecil dari nilai α (0,05) yang berarti ada perbedaan antar kelompok yang terbentuk.

2.11 Konsumsi Bahan Bakar Minyak (BBM)

Bahan Bakar Minyak (BBM) adalah bahan bakar mineral cair yang diperoleh dari hasil tambang pengeboran sumur-sumur minyak mentah atau *crude oil* yang selanjutnya dilakukan pengolahan lebih lanjut sampai menghasilkan bermacam-macam bahan bakar yang memiliki kualitas yang berbeda-beda (Koesoemadinata, 1980). Di Indonesia, Bahan Bakar Minyak (BBM) diproduksi oleh Badan Usaha Milik Negara (BUMN) yang bergerak di bidang energi minyak dan gas bumi yaitu PT. Pertamina. Produk-produk Bahan Bakar Minyak (BBM) yang diproduksi oleh PT. Pertamina diantaranya: Premium, Minyak Tanah, Minyak Solar, *Aviation Gasoline* (BBM pesawat udara), *Aviation Turbine Fuel* (BBM pesawat udara berturbin), Bio Solar, Pertamina, Pertamina Plus, Pertamina Dex, Peralite, dan Bio Pertamina. Beberapa jenis produk BBM tersebut didistribusikan oleh PT. Pertamina ke Stasiun Pengisian Bahan Bakar Umum (SPBU) yang selanjutnya dijual untuk memenuhi kebutuhan konsumsi BBM di masyarakat.

Jumlah konsumsi Bahan Bakar Minyak (BBM) di masyarakat dipengaruhi oleh kondisi tipologi dan sistem transportasi pada suatu wilayah. Kondisi tipologi suatu wilayah dapat ditinjau dari beberapa faktor diantaranya: kepadatan penduduk, jumlah penduduk, tataguna lahan, dan PDRB. Sedangkan kondisi sistem transportasi dapat ditinjau berdasarkan panjang jalan, pola jaringan jalan, kondisi jalan, angkutan umum penumpang, angkutan barang, panjang trayek angkutan umum serta kendaraan pribadi (Handayani, 2012). Jumlah konsumsi Bahan Bakar Minyak di masyarakat juga dipengaruhi oleh kondisi kemacetan lalu lintas yang ada di daerah tersebut, hasil penelitian yang dilakukan oleh Alamsyah & Anshori (2004) menyatakan bahwa jika lalu lintas di suatu wilayah terjadi kemacetan maka konsumsi BBM akan bertambah

(Halaman ini sengaja dikosongkan)

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Penelitian ini menggunakan data sekunder tentang jumlah konsumsi Bahan Bakar Minyak (BBM) dan faktor-faktor yang mempengaruhi perbedaan tingkat konsumsi BBM pada setiap Kabupaten/Kota di Jawa Timur. Data diambil dari laporan realisasi Bahan Bakar Minyak Wilayah Jawa Timur yang dikeluarkan oleh PT. Pertamina (Persero) MOR V Surabaya, laporan tahunan tentang jumlah transportasi di Kabupaten/Kota yang dikeluarkan oleh Badan Pendapatan Daerah Provinsi Jawa Timur, dan publikasi laporan tahunan tentang kependudukan, ekonomi, dan infrastruktur yang dikeluarkan oleh Badan Pusat Statistik (BPS) Provinsi Jawa Timur. Data tersebut merupakan data tahun 2011-2015 dengan unit pengamatan yang diambil pada tingkat Kabupaten/ Kota di Provinsi Jawa Timur sebanyak 38 Kabupaten/ Kota, namun dalam penelitian ini dilakukan penggabungan data (*data merging*) pada beberapa Kabupaten/Kota sehingga jumlah Kabupaten/Kota di Jawa Timur yang diambil sebagai unit pengamatan menjadi 31 Kabupaten/Kota.

3.2 Variabel Penelitian

Variabel penelitian yang digunakan dalam penelitian ini terdiri dari 16 variabel prediktor (X).

Tabel 3.1 Variabel Penelitian

Variabel	Definisi Operasional	Tipe Variabel	Satuan
Jumlah konsumsi BBM jenis Premium (X_1)	Jumlah konsumsi Bahan Bakar Minyak (BBM) jenis Premium di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011- 2015.	Rasio (Kontinyu)	Kiloliter

Tabel 3.1 Variabel Penelitian (Lanjutan)

Jumlah konsumsi BBM jenis Pertamina (X_2)	Jumlah konsumsi Bahan Bakar Minyak (BBM) jenis Pertamina di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011-2015.	Rasio (Kontinyu)	Kiloliter
Jumlah konsumsi BBM jenis Pertamina Plus (X_3)	Jumlah konsumsi Bahan Bakar Minyak (BBM) jenis Pertamina Plus di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011-2015.	Rasio (Kontinyu)	Kiloliter
Jumlah konsumsi BBM jenis Solar (X_4)	Jumlah konsumsi Bahan Bakar Minyak (BBM) jenis Solar di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011-2015.	Rasio (Kontinyu)	Kiloliter
Jumlah konsumsi BBM jenis Pertamina Dex (X_5)	Jumlah konsumsi Bahan Bakar Minyak (BBM) jenis Pertamina Dex di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011-2015.	Rasio (Kontinyu)	Kiloliter
Jumlah kendaraan sedan (X_6)	Jumlah kendaraan sedan di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011-2015.	Rasio (Diskrit)	Unit
Jumlah kendaraan jeep (X_7)	Jumlah kendaraan jeep di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011-2015.	Rasio (Diskrit)	Unit
Jumlah kendaraan <i>station wagon</i> (X_8)	Jumlah kendaraan <i>station wagon</i> di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011-2015.	Rasio (Diskrit)	Unit

Tabel 3.1 Variabel Penelitian (Lanjutan)

Jumlah kendaraan <i>bus</i> (X_9)	Jumlah kendaraan <i>bus</i> di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011-2015.	Rasio (Diskrit)	Unit
Jumlah kendaraan <i>truck</i> (X_{10})	Jumlah kendaraan <i>truck</i> di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011-2015.	Rasio (Diskrit)	Unit
Jumlah kendaraan sepeda motor (X_{11})	Jumlah kendaraan sepeda motor di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011-2015.	Rasio (Diskrit)	Unit
Jumlah kendaraan alat berat (X_{12})	Jumlah kendaraan alat berat di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011-2015.	Rasio (Diskrit)	Unit
Jumlah penduduk (X_{13})	Jumlah penduduk di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011-2015.	Rasio (Diskrit)	Orang
Produk Domestik Regional Bruto Harga Berlaku (X_{14})	Produk Domestik Regional Bruto (PDRB) atas dasar harga berlaku di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011-2015	Rasio (Kontinyu)	Miliar Rupiah
Jumlah perahu motor (X_{15})	Jumlah perahu motor di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011-2015.	Rasio (Diskrit)	Unit
Jumlah kapal motor (X_{16})	Jumlah kapal motor di Kabupaten/Kota di Provinsi Jawa Timur tahun 2011-2015.	Rasio (Diskrit)	Unit

Berdasarkan 16 variabel penelitian seperti yang ditampilkan pada Tabel 3.1, akan dibagi menjadi beberapa *subset* data yang merupakan kombinasi dari beberapa variabel penelitian. Pembentukan *subset* data berdasarkan pada karakteristik pengelompokan yang ingin diteliti dan disesuaikan dengan kebutuhan analisis penelitian. *Subset* data yang terbentuk akan digunakan sebagai dasar pengelompokan Kabupaten/Kota di Jawa Timur. *Subset* data yang dibentuk dalam penelitian ini adalah sebagai berikut.

Tabel 3.2 Kombinasi Variabel pada *Subset* Data Penelitian

<i>Subset</i> Data	Variabel Penelitian	Karakteristik Pengelompokan
Subset 1 (s_1)	$X_1, X_{11}, X_{13}, X_{14}, X_{15}$	Premium, Sepeda motor, Penduduk, PDRB, dan Perahu motor
Subset 2 (s_2)	$X_4, X_9, X_{10}, X_{13}, X_{14}, X_{16}$	Solar, <i>Bus</i> , <i>Truck</i> , Penduduk, PDRB, dan Kapal motor
Subset 3 (s_3)	$X_1, X_2, X_3, X_6, X_7, X_8, X_{11}, X_{13}, X_{14}, X_{15}$	Premium, Pertamina, Pertamina Plus, Sedan, <i>Jeep</i> , <i>Station Wagon</i> , Sepeda motor, Penduduk, PDRB, dan Perahu motor
Subset 4 (s_4)	$X_4, X_5, X_9, X_{10}, X_{12}, X_{13}, X_{14}, X_{16}$	Solar, Pertamina Dex, <i>Bus</i> , <i>Truck</i> , alat berat Penduduk, PDRB, dan Kapal motor

Berdasarkan Tabel 3.2, setiap *subset* data yang terbentuk akan dibedakan untuk setiap tahun yaitu mulai tahun 2011 sampai tahun 2015, sehingga jumlah *subset* data yang terbentuk sebanyak 20 *subset* data. Pada penelitian ini dilakukan analisis pengelompokan (*clustering*) Kabupaten/Kota di Jawa Timur dengan menggunakan metode *model-based clustering* berdasarkan kombinasi variabel yang terdapat pada setiap *subset* data. Struktur data yang terbentuk untuk setiap *subset* data pada setiap tahun ditunjukkan pada Tabel 3.3 sebagai berikut.

Tabel 3.3 Struktur *Subset* Data Penelitian Setiap Tahun

Kode Wilayah	X_1	X_2	X_3	X_4	X_5	X_6	...	X_p	Prediksi Kelompok (\hat{Z})
1	$x_{1.1}$	$x_{2.1}$	$x_{3.1}$	$x_{4.1}$	$x_{5.1}$	$x_{6.1}$...	$x_{p.1}$	\hat{Z}_1
2	$x_{1.2}$	$x_{2.2}$	$x_{3.2}$	$x_{4.2}$	$x_{5.2}$	$x_{6.2}$...	$x_{p.2}$	\hat{Z}_2
3	$x_{1.3}$	$x_{2.3}$	$x_{3.3}$	$x_{4.3}$	$x_{5.3}$	$x_{6.3}$...	$x_{p.3}$	\hat{Z}_3
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
31	$x_{1.31}$	$x_{2.31}$	$x_{3.31}$	$x_{4.31}$	$x_{5.31}$	$x_{6.31}$...	$x_{p.31}$	\hat{Z}_{31}

Sedangkan kode wilayah untuk wilayah Kabupaten/Kota dengan urutan seperti pada Tabel 3.3 dapat dilihat pada Tabel 3.4 sebagai berikut.

Tabel 3.4 Daftar Kabupaten/Kota di Provinsi Jawa Timur

Kode Wilayah	Wilayah	Kode Wilayah	Wilayah	Kode Wilayah	Wilayah
1.	Kab. Pacitan	7.	Kab. Malang & Batu	13.	Probolinggo (Kab-Kota)
2.	Kab. Ponorogo	8.	Kab. Lumajang	14.	Pasuruan (Kab-Kota)
3.	Kab. Trenggalek	9.	Kab. Jember	15.	Kab. Sidoarjo
4.	Kab. Tulungagung	10.	Kab. Banyuwangi	16.	Mojokerto (Kab-Kota)
5.	Blitar (Kab-Kota)	11.	Kab. Bondowoso	17.	Kab. Jombang
6.	Kediri (Kab-Kota)	12.	Kab. Situbondo	18.	Kab. Nganjuk

Tabel 3.4 Daftar Kabupaten/Kota di Provinsi Jawa Timur
(Lanjutan)

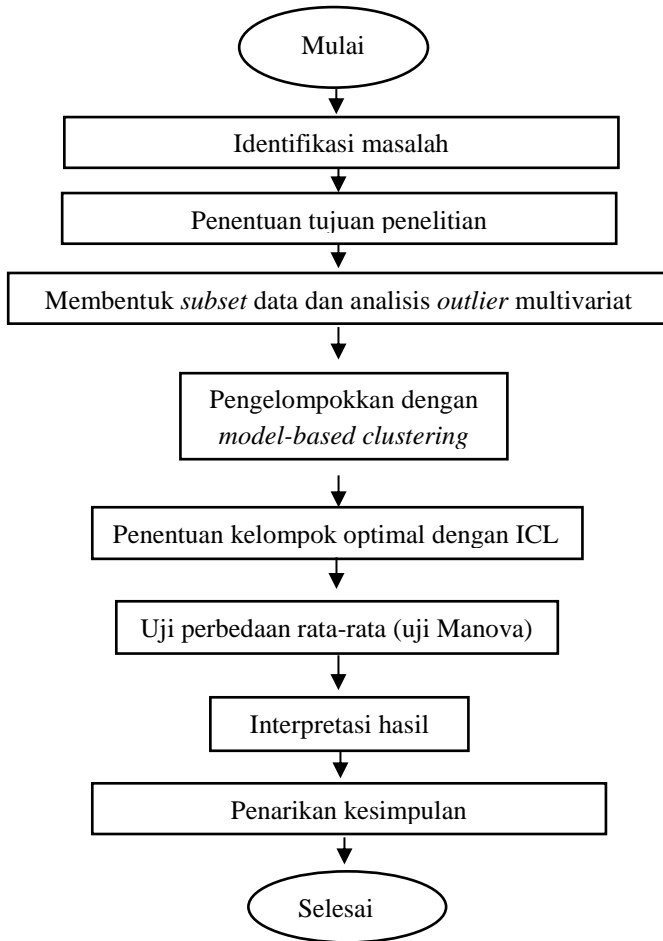
Kode Wilayah	Wilayah	Kode Wilayah	Wilayah	Kode Wilayah	Wilayah
19.	Madiun (Kab-Kota)	24.	Kab. Lamongan	29.	Kab. Sumenep
20.	Kab. Magetan	25.	Kab. Gresik	30	Kota Malang
21.	Kab. Ngawi	26.	Kab. Bangkalan	31	Kota Surabaya
22.	Kab. Bojonegoro	27.	Kab. Sampang		
23.	Kab. Tuban	28.	Kab. Pamekasan		

3.3 Langkah-langkah Penelitian

Langkah-langkah untuk melakukan pemetaan tingkat konsumsi BBM dengan *model-based clustering* adalah sebagai berikut:

1. Membuat *subset* data dengan mengkombinasikan beberapa variabel penelitian.
2. Melakukan pengujian *outlier* multivariat terhadap setiap *subset* data penelitian.
3. Melakukan pengujian distribusi *t* multivariat terhadap setiap *subset* data penelitian.
4. Melakukan pengelompokkan objek penelitian dengan metode *model-based clustering*.
5. Menentukan jumlah kelompok optimal dengan kriteria *Integrated Completed Likelihood* (ICL).
6. Melakukan uji perbedaan rata-rata terhadap hasil pengelompokkan dengan menggunakan uji Manova.
7. Melakukan interpretasi hasil *clustering* yang didapatkan.
8. Membuat kesimpulan berdasarkan hasil analisis.

Secara umum, langkah analisis disajikan pada Gambar 3.1 sebagai berikut.



Gambar 3.1 Diagram Alir Penelitian

(Halaman ini sengaja dikosongkan)

BAB IV

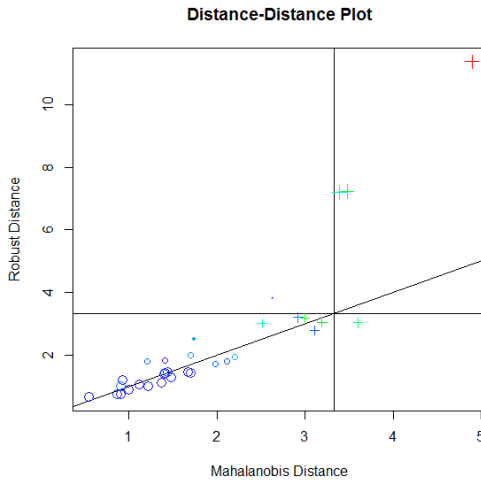
HASIL DAN PEMBAHASAN

Pada bab ini akan dijelaskan beberapa hal yang dilakukan untuk mencapai tujuan menjawab permasalahan dalam penelitian ini. Hal-hal yang akan dibahas diantaranya meliputi deteksi *outlier* multivariat, pengelompokkan Kabupaten/Kota di Jawa Timur menggunakan *model-based clustering* dengan kriteria *Integrated Completed Likelihood* (ICL) dan analisis karakteristik tingkat konsumsi BBM di Kabupaten/Kota di Jawa Timur. Tujuan utama penelitian ini menggunakan metode *model-based clustering* adalah mengetahui Kabupaten/Kota di Jawa Timur yang memiliki kemiripan karakteristik berdasarkan hasil pengelompokkan (*clustering*) menurut tingkat konsumsi Bahan Bakar Minyak (BBM) serta faktor-faktor yang mempengaruhi perbedaan tingkat konsumsi BBM.

4.1 Deteksi *Outlier* Multivariat

Deteksi *outlier* multivariat dilakukan untuk membuktikan bahwa data tingkat konsumsi Bahan Bakar Minyak (BBM) pada Kabupaten/Kota di Jawa Timur cenderung memuat *outlier*. Pada penelitian ini dilakukan deteksi *outlier* pada masing-masing *subset* data tingkat konsumsi BBM yaitu *subset* data s_1 , s_2 , s_3 , dan s_4 dari tahun 2011 sampai tahun 2015 . Proses deteksi *outlier* multivariat pada data penelitian dilakukan dengan menghitung jarak mahalanobis dan jarak *robust* untuk semua objek pengamatan. Selanjutnya dibandingkan dengan kuantil dari distribusi $\chi^2_{p;0,975}$ sebagai nilai *cut off*. Nilai *cut off* pada masing-masing *subset* data tingkat konsumsi BBM bisa berbeda-beda, tergantung dari jumlah variabel (p) dari *subset* data tersebut. Suatu objek pengamatan dikatakan *outlier* jika nilai jarak mahalanobis dan jarak *robust*

lebih besar dari nilai $\sqrt{\chi^2_{p;0,975}}$.



Gambar 4.1 *Plot Jarak Mahalanobis terhadap Jarak Robust Subset Data S_1 2015*

Plot jarak mahalanobis terhadap jarak robust *subset* data s_1 2015 seperti yang ditunjukkan oleh Gambar 4.1, dilakukan dengan *software R package mvoutlier* menggunakan fungsi *distance-distance plot (dd plot)*. Data tidak terdeteksi *outlier* jika titik-titik pada *plot* terletak di sekitar garis lurus, perpotongan jarak mahalanobis dan jarak robust. Berdasarkan hasil *plot* pada Gambar 4.1 menunjukkan terdapat tiga Kabupaten/Kota yang terdeteksi sebagai *outlier* pada *subset* data tingkat konsumsi BBM S_1 2015. Ketiga Kabupaten/Kota tersebut adalah Kabupaten Malang & Batu, Kabupaten Jember, dan Kota Surabaya. Selain deteksi *outlier* multivariat melalui *plot*, *outlier* dapat dideteksi melalui *output* nilai jarak robust dan mahalanobis dari setiap Kabupaten/Kota di Jawa Timur sebagaimana terlampir (Lampiran 6). Keterangan “TRUE” berarti Kabupaten/Kota terdeteksi sebagai *outlier* pada *subset* data s_1 2015. Sebaliknya, keterangan “FALSE” menunjukkan bahwa Kabupaten/Kota tidak terdeteksi sebagai *outlier*.

Tabel 4.1 Jumlah Kabupaten/Kota Terdeteksi Sebagai *Outlier* pada Kuantil 90 Persen

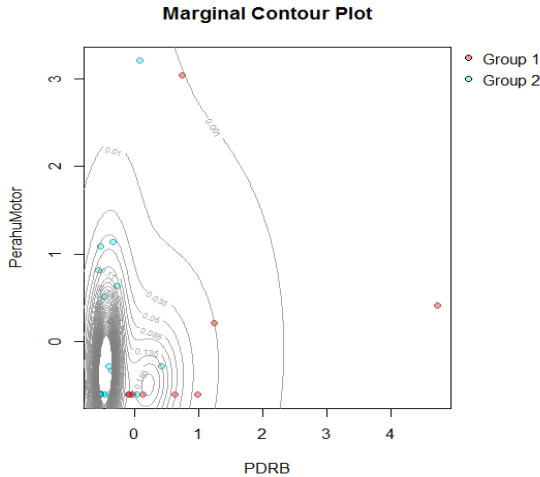
Subset Data	Tahun				
	2011	2012	2013	2014	2015
s₁	3	3	3	3	3
s₂	3	3	3	3	3
s₃	3	3	3	2	2
s₄	3	3	3	3	3

Berdasarkan Tabel 4.1 menunjukkan bahwa semua *subset* data tingkat konsumsi BBM terdapat beberapa Kabupaten/Kota yang terdeteksi sebagai *outlier* pada kuantil 90 persen. Berdasarkan *output* hasil deteksi *outlier* dengan *package mvoutlier* (Lampiran 7 sampai dengan Lampiran 10), Kota Surabaya menjadi menjadi daerah yang selalu terdeteksi sebagai *outlier* pada semua *subset* data dari tahun 2011 sampai tahun 2015. Hal tersebut disebabkan data jumlah konsumsi Bahan Bakar Minyak di kota Surabaya paling tinggi dibandingkan dengan Kabupaten/Kota lainnya di Jawa Timur. Selain itu data beberapa variabel lainnya seperti jumlah sepeda motor, sedan, *jeep*, *station wagon*, *bus*, *truck*, penduduk, dan Produk Domestik Regional Bruto (PDRB) dibandingkan dengan Kabupaten/Kota lainnya di Jawa Timur.

4.2 Pengelompokkan Menggunakan *Model-Based Clustering* dengan Kriteria *Integrated Completed Likelihood* (MBC-ICL)

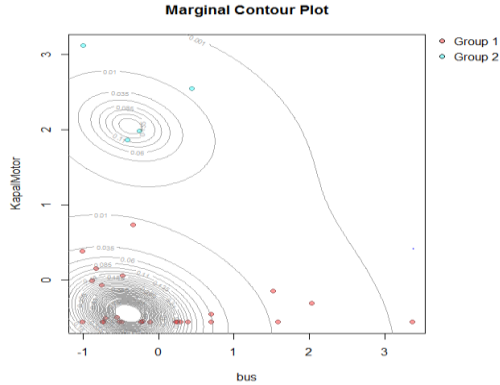
Pada penelitian ini, semua *subset* data yang terbentuk diasumsikan memenuhi distribusi *t* multivariat sehingga dapat dilakukan pengelompokkan dengan menggunakan metode *model-based clustering* melalui *package teigen*. Pemilihan model terbaik dan jumlah kelompok optimal yang terbentuk dapat ditentukan melalui nilai ICL terbesar. Berdasarkan *ouput* hasil pengelompokkan *subset* data *s₁2015* seperti pada Lampiran 31 menunjukkan bahwa jumlah kelompok optimal yang terbentuk sebanyak dua kelompok dengan model terbaik pada pengelompokkannya adalah CUCC. Model terbaik CUCC tersebut

menunjukkan bahwa model matriks varians kovarians kedua kelompok adalah $\Sigma_g = \lambda \mathbf{D}_g \mathbf{A} \mathbf{D}^T$. Ini berarti kedua kelompok memiliki volume elips, kontur fungsi kepadatan, dan derajat bebas yang sama, namun memiliki orientasi yang berbeda (Lampiran 32). Hasil *marginal contour plot* yang terbentuk pada pengelompokkan *subset* data $s_1 2015$ adalah sebagai berikut.



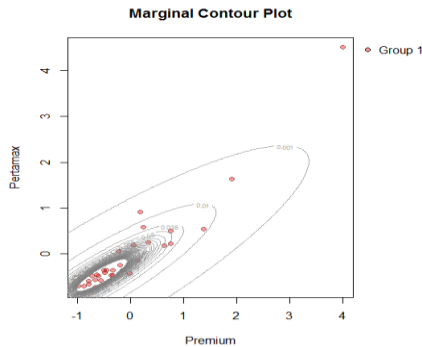
Gambar 4.2 *Marginal Contour Plot Subset Data S₁ 2015*

Hasil yang sama untuk jumlah kelompok optimal ditunjukkan pada pengelompokkan *subset* data $s_2 2015$ dengan terbentuk dua kelompok namun memiliki model pengelompokkan yang berbeda yaitu CCCC, yang menunjukkan bahwa model matriks varians kovarians kedua kelompok adalah $\Sigma_g = \Sigma = \lambda \mathbf{D} \mathbf{A} \mathbf{D}^T$. Ini berarti kedua kelompok memiliki volume elips, orientasi, kontur fungsi kepadatan, dan derajat bebas yang sama, sehingga menyebabkan matriks varians dan kovarians dari kedua kelompok tersebut menjadi sama. Hasil *marginal contour plot* yang terbentuk pada pengelompokkan *subset* data $s_2 2015$ adalah sebagai berikut.

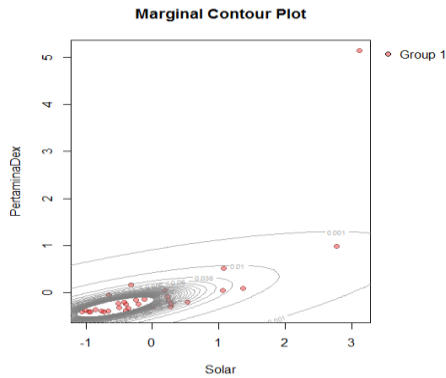


Gambar 4.3 *Marginal Contour Plot Subset Data S₂ 2015*

Sedangkan hasil yang berbeda untuk jumlah kelompok optimal ditunjukkan pada pengelompokkan *subset* data s_3 2015 dan s_4 2015 dengan terbentuk hanya satu kelompok dengan model pengelompokkannya yaitu UUUU, yang menunjukkan bahwa model matriks varians kovarians kedua kelompok adalah $\Sigma_g = \lambda_g \mathbf{D}_g \mathbf{A}_g \mathbf{D}_g^T$. Ini berarti kedua kelompok memiliki volume elips, orientasi, kontur fungsi kepadatan, dan derajat bebas yang berbeda, sehingga menyebabkan matriks varians dan kovarians dari kedua kelompok tersebut menjadi berbeda. Hasil *marginal contour plot* yang terbentuk pada pengelompokkan *subset* data s_3 2015 dan s_4 2015 adalah sebagai berikut.



Gambar 4.4 *Marginal Contour Plot Subset Data S₃ 2015*



Gambar 4.5 *Marginal Contour Plot Subset Data S₄ 2015*

Dengan cara yang sama, dapat diperoleh pengelompokan untuk *subset-subset* data yang lain. Hasil jumlah kelompok optimal yang terbentuk pada setiap *subset* data dapat ditampilkan sebagai berikut.

Tabel 4.2 Jumlah Kelompok Optimal Metode MBC-ICL

Subset Data	Tahun				
	2011	2012	2013	2014	2015
s ₁	2	2	2	2	2
s ₂	1	2	2	2	2
s ₃	1	1	1	1	1
s ₄	1	1	1	1	1

Jumlah kelompok optimal pada setiap *subset* data seperti yang ditampilkan pada Tabel 4.2, dapat diketahui bahwa terdapat 11 *subset* data yang menghasilkan jumlah kelompok optimal sebanyak satu kelompok. Hal ini menunjukkan bahwa karakteristik konsumsi BBM Kabupaten/Kota di Jawa Timur pada *subset-subset* data tersebut adalah sama. Sedangkan pada 9 *subset* data lainnya menghasilkan jumlah kelompok optimal sebanyak dua kelompok yang menunjukkan bahwa karakteristik konsumsi BBM Kabupaten/Kota di Jawa Timur pada *subset-subset* data tersebut berbeda pada setiap kelompoknya, namun hal ini perlu dilakukan pengujian lebih lanjut untuk mengetahui apakah antar kelompok

memiliki karakteristik yang berbeda secara signifikan. Salah satu pengujian yang dapat dilakukan adalah uji beda rata-rata melalui uji Manova.

4.3 Uji Kesamaan Mean Antar Kelompok

Pada penelitian ini dilakukan uji kesamaan *mean* antar kelompok dengan tujuan untuk mengetahui apakah antar kelompok yang terbentuk memang berbeda secara signifikan, yaitu dengan menggunakan uji beda *mean* melalui uji Manova. Manova merupakan salah satu analisis multivariat yang digunakan untuk mengetahui apakah vektor rata-rata populasi sama. Hipotesis yang digunakan untuk uji beda rata-rata antar dua kelompok adalah sebagai berikut.

$H_0 : \tau_1 = \tau_2$ (tidak ada perbedaan vektor *mean* antar kelompok)

$H_1 : \tau_1 \neq \tau_2$ (ada perbedaan vektor *mean* antar kelompok)

Keputusan tolak jika *p_value* lebih kecil dari nilai α (0,05) yang berarti kelompok 1 berbeda secara signifikan dengan kelompok 2. Pengujian ini akan dilakukan untuk 9 *subset* data yang memiliki jumlah kelompok optimal sebanyak 2 kelompok. Hasil uji kesamaan kelompok pada masing-masing *subset* data tersebut dapat ditampilkan sebagai berikut.

Tabel 4.3 Hasil Uji Kesamaan Kelompok (Uji Manova) dengan Statistik Uji *Wald-Type Test*

Subset	Nilai Wald-Type Test	<i>p_value</i>
s ₁ 2011	1635,239	0,000
s ₁ 2012	1623,038	0,000
s ₁ 2013	1568,316	0,000
s ₁ 2014	24,168	$2,02 \times 10^{-4}$
s ₁ 2015	31,087	$9,00 \times 10^{-6}$
s ₂ 2012	871,343	0,000
s ₂ 2013	1040,001	0,000
s ₂ 2014	1187,071	0,000
s ₂ 2015	1103,392	0,000

Berdasarkan hasil uji kesamaan kelompok yang dilakukan melalui *package* MANOVA.RM seperti yang ditampilkan pada Tabel 4.3, menunjukkan bahwa statistik *Wald-Type Test* memiliki *p_value* lebih kecil dari nilai α (0,05) untuk setiap *subset* data. Hal ini menunjukkan bahwa adanya perbedaan vektor *mean* antar kelompok pada setiap *subset* data, dimana setiap *subset* data yang diuji masing-masing terbentuk 2 kelompok sehingga berdasarkan hasil pengujian dapat disimpulkan bahwa kelompok 1 berbeda secara signifikan dengan kelompok 2. Selanjutnya perlu dilakukan analisis pada masing-masing *subset* data untuk mengetahui karakteristik antar kelompok yang terbentuk.

4.4 Analisis Karakteristik Tingkat Konsumsi BBM di Jawa Timur

Setelah dilakukan pengujian kesamaan kelompok yang menunjukkan adanya perbedaan vektor *mean* antar kelompok, maka perlu dilakukan analisis untuk mengetahui karakteristik antar kelompok yang terbentuk. Berikut ini adalah hasil analisis karakteristik antar kelompok untuk *subset* data s_1 tahun 2011 sampai tahun 2015 dan *subset* data s_2 tahun 2012 sampai tahun 2015.

4.4.1 Karakteristik Konsumsi BBM di Jawa Timur Berdasarkan Jumlah Sepeda Motor, Perahu Motor Penduduk, dan PDRB (s_1)

Konsumsi BBM pada *subset* data s_1 adalah BBM jenis Premium. Pada penelitian ini, dilakukan analisis untuk mengetahui pola sebaran pengelompokkan dan karakteristik konsumsi BBM jenis Premium berdasarkan jumlah sepeda motor, perahu motor, penduduk, dan PDRB Kabupaten/Kota di Jawa Timur pada periode tahun 2011 sampai tahun 2015. Peta sebaran pengelompokkan *subset* data s_1 pada tahun 2011 sampai tahun 2015 ditampilkan sebagai berikut.



Gambar 4.9 Peta Sebaran Pengelompokkan *Subset* Data (s_1) Tahun 2014



Gambar 4.10 Peta Sebaran Pengelompokkan *Subset* Data (s_1) Tahun 2015

Berdasarkan peta sebaran pengelompokkan *subset* data (s_1) tahun 2011 sampai tahun 2015 menunjukkan bahwa setiap tahun jumlah kelompok yang terbentuk selalu sama yaitu sebanyak 2 kelompok. Sebanyak lebih dari 50 persen Kabupaten/Kota di Jawa Timur konsisten berada di kelompok 1 selama periode tahun 2011 sampai tahun 2015. Sedangkan sebanyak 2 Kabupaten/Kota yang konsisten berada di kelompok 2, yaitu Kabupaten Malang & Batu dan Kabupaten Jember. Pada periode tahun 2011 sampai tahun 2013 memiliki pola sebaran pengelompokkan yang konsisten sama dengan terbentuknya 2 kelompok. Masing-masing kelompok

tersebut juga mempunyai anggota yang sama pada setiap tahunnya, yaitu sebanyak 2 Kabupaten/Kota (Kabupaten Malang & Batu dan Kabupaten Jember) berada di kelompok 2, sedangkan 29 Kabupaten/Kota lainnya berada di kelompok 1 (Lampiran 39). Hasil berbeda ditunjukkan oleh sebaran pengelompokkan pada tahun 2014 dan tahun 2015. Jika dibandingkan dengan tahun sebelumnya, sebaran pengelompokkan pada tahun 2014 menunjukkan adanya peningkatan jumlah keanggotaan pada kelompok 2 yaitu menjadi 10 Kabupaten/Kota. Berdasarkan Lampiran 39 menunjukkan bahwa pada tahun 2014, Kabupaten Malang & Batu dan Kabupaten Jember konsisten berada di kelompok 2 dan terdapat penambahan keanggotaan pada kelompok tersebut sebanyak 8 Kabupaten/Kota, yaitu Kabupaten Banyuwangi, Kabupaten Sidoarjo, Kabupaten Gresik, Kota Malang, Kota Surabaya, Kediri, Pasuruan, dan Mojokerto. Sedangkan pada tahun 2015, jumlah keanggotaan masing-masing kelompok sama dengan tahun 2014, yaitu sebanyak 10 Kabupaten/Kota berada di kelompok 2 dan 21 Kabupaten/Kota lainnya berada di kelompok 1, namun terjadi perubahan keanggotaan pada masing-masing kelompok. Berdasarkan perbandingan hasil pengelompokkan sebagaimana terlampir (Lampiran 39) menunjukkan secara keseluruhan pada periode tahun 2011 sampai tahun 2015, terdapat sebanyak 21 Kabupaten/Kota yang konsisten tidak mengalami perubahan keanggotaan kelompok, hal ini dapat dilihat pada Kabupaten Malang & Batu dan Kabupaten Jember yang konsisten berada di kelompok 2, sedangkan 10 Kabupaten/Kota lainnya cenderung berubah-ubah dalam keanggotaan kelompok.

Tabel 4.4 Nilai *Mean* pada Hasil Pengelompokkan *Subset* Data (s₁) Tahun 2011-2015

Variabel		Nilai <i>Mean Subset</i> Data s₁				
		2011	2012	2013	2014	2015
Premium	1	107454	119120	123273	80369	82216
	2	205511	229232	238572	243863	200650

Tabel 4.4 Nilai *Mean* pada Hasil Pengelompokkan *Subset* Data (s_1) Tahun 2011-2015 (Lanjutan)

Variabel		Nilai <i>Mean</i> <i>Subset</i> Data s_1				
		2011	2012	2013	2014	2015
Sepeda Motor	1	282130	315716	355922	269507	267095
	2	448884	510017	561780	716506	751558
Penduduk	1	1131822	1139777	1147480	958159	1009508
	2	2508914	2526530	2543144	1848886	1764789
PDRB	1	35530	39640	44107	21728	26715
	2	45725	51102	56405	109085	113336
Perahu Motor	1	1114	1030	947	661	1100
	2	1001	598	194	1248	299

Berdasarkan Tabel 4.4, hasil pengelompokkan *subset* data s_1 pada tahun 2015 menunjukkan bahwa kelompok 1 merupakan kelompok Kabupaten/Kota di Jawa Timur dengan karakteristik konsumsi BBM jenis Premium yang rendah (82.216 kiloliter), jumlah sepeda motor yang rendah (267.095 unit), jumlah penduduk yang rendah (1,009 juta jiwa), dan PDRB yang rendah (26.715 miliar dollar). Sedangkan pada kelompok 2 merupakan Kabupaten/Kota dengan karakteristik konsumsi BBM jenis Premium yang tinggi (200.650 kiloliter), jumlah sepeda motor yang tinggi (751.558 unit), jumlah penduduk yang tinggi (1,76 juta jiwa), dan PDRB yang tinggi (113.336 miliar dollar). Hasil karakteristik yang cenderung sama juga didapatkan untuk pengelompokkan *subset* data pada tahun-tahun sebelumnya.

4.4.2 Karakteristik Konsumsi BBM Jawa Timur Berdasarkan Jumlah Bus, Truck, Kapal, Penduduk, dan PDRB (s_2)

Konsumsi BBM pada *subset* data s_2 adalah BBM jenis Solar. Pada penelitian ini, dilakukan analisis untuk mengetahui pola sebaran pengelompokkan dan karakteristik konsumsi BBM jenis Solar berdasarkan jumlah *bus*, *truck*, kapal, penduduk, dan PDRB

Kabupaten/Kota di Jawa Timur pada periode tahun 2012 sampai tahun 2015. Peta sebaran pengelompokkan *subset* data s_2 pada tahun 2012 sampai tahun 2015 ditampilkan sebagai berikut.



Gambar 4.11 Peta Sebaran Pengelompokkan *Subset* Data (s_2) Tahun 2012



Gambar 4.12 Peta Sebaran Pengelompokkan *Subset* Data (s_2) Tahun 2013



Gambar 4.13 Peta Sebaran Pengelompokkan *Subset* Data (s_2) Tahun 2014



Gambar 4.14 Peta Sebaran Pengelompokkan *Subset* Data (s_2) Tahun 2015

Berdasarkan peta sebaran pengelompokkan *subset* data (s_2) tahun 2012 sampai tahun 2015 menunjukkan bahwa setiap tahun jumlah kelompok yang terbentuk selalu sama yaitu sebanyak 2 kelompok. Masing-masing kelompok tersebut juga mempunyai anggota yang sama pada setiap tahunnya, yaitu sebanyak 2 Kabupaten/Kota (Pasuruan dan Kabupaten Sidoarjo) berada di kelompok 1, sedangkan 29 Kabupaten/Kota lainnya berada di kelompok 2 (Lampiran 40). Hal ini berarti bahwa selama periode tahun 2012 sampai tahun 2015, tidak terjadi perubahan pengelompokkan ataupun keanggotaan kelompok pada setiap

tahunnya sehingga dapat dikatakan bahwa hasil pengelompokkan antar tahun pada periode tersebut adalah sama.

Tabel 4.5 Nilai *Mean* pada Hasil Pengelompokkan *Subset* Data (s_2) Tahun 2012-2015

Variabel		Nilai <i>Mean Subset</i> Data s_2			
		2012	2013	2014	2015
Solar	1	173107	179803	182502	161239
	2	59880	60990	59928	55410
Bus	1	1250	1363	1537	1600
	2	591	645	704	763
Truck	1	20735	23228	26345	28665
	2	13019	14507	15924	17107
Penduduk	1	1875444	1898991	1923380	1946941
	2	1184679	1191904	1198739	1205299
PDRB	1	93955	103760	115948	128304
	2	36684	40841	45354	49578
Kapal	1	48	43	64	85
	2	742	879	934	975

Hasil pengelompokkan *subset* data s_2 pada tahun 2015 seperti yang ditampilkan pada Tabel 4.5 menunjukkan bahwa kelompok 1 merupakan kelompok Kabupaten/Kota di Jawa Timur dengan karakteristik konsumsi BBM jenis Solar yang tinggi (161.239 kiloliter), jumlah *bus* yang tinggi (1600 unit), jumlah *truck* yang tinggi (28.665 unit), jumlah penduduk yang tinggi (1,95 juta jiwa), dan PDRB yang tinggi (128.304 miliar dollar). Sedangkan pada kelompok 2 merupakan Kabupaten/Kota dengan karakteristik konsumsi BBM jenis Solar yang rendah (55.410 kiloliter), jumlah *bus* yang rendah (763 unit), jumlah *truck* yang rendah (17.107 unit), jumlah penduduk yang rendah (1,2 juta jiwa), dan PDRB yang rendah (49.578 miliar dollar). Secara keseluruhan, hasil pengelompokkan pada tahun 2012 sampai tahun 2015 mempunyai karakteristik yang sama pada setiap tahunnya.

(Halaman ini sengaja dikosongkan)

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan, dapat diambil kesimpulan sebagai berikut:

1. Berdasarkan peta sebaran pengelompokkan *subset* data (s_1) tahun 2011 sampai tahun 2015 menunjukkan bahwa sebagian besar wilayah Kabupaten/Kota di Jawa Timur konsisten berada di kelompok 1. Sedangkan berdasarkan peta sebaran pengelompokkan *subset* data (s_2) tahun 2012 sampai tahun 2015 menunjukkan bahwa setiap kelompok yang terbentuk mempunyai anggota yang sama pada setiap tahunnya, dengan sebagian besar wilayah Kabupaten/Kota di Jawa Timur konsisten berada di kelompok 2.
2. Hasil pengelompokkan dengan *model-based clustering* dengan menggunakan kriteria *Integrated Completed Likelihood* (ICL) menunjukkan bahwa jumlah kelompok optimal yang terbentuk pada *subset* data s_1 dan s_2 selama periode 2012-2015 selalu konsisten sama dengan terbentuknya 2 kelompok.
3. Berdasarkan hasil pengelompokkan pada *subset* data s_1 tahun 2011-2015 menunjukkan lebih dari 50 persen Kabupaten/ Kota di Jawa Timur konsisten berada di kelompok 1, yaitu kelompok yang memiliki karakteristik konsumsi BBM jenis Premium, jumlah sepeda motor, penduduk, dan PDRB yang tergolong rendah. Secara keseluruhan, karakteristik hasil pengelompokkan *subset* data s_1 tahun 2011-2015 cenderung sama pada setiap tahun. Sedangkan hasil pengelompokkan pada *subset* data s_2 tahun 2012-2015 menunjukkan 29 Kabupaten/ Kota di Jawa Timur konsisten berada di kelompok 2, yaitu kelompok yang memiliki karakteristik konsumsi BBM jenis Solar, jumlah *bus*, *truck*, penduduk, PDRB, dan kapal motor yang tergolong rendah. Secara keseluruhan,

karakteristik hasil pengelompokkan *subset* data s_2 tahun 2012-2015 konsisten sama pada setiap tahun.

5.2 Saran

Saran yang diberikan berdasarkan hasil analisis yang telah dilakukan sebelumnya adalah sebagai berikut.

1. Berdasarkan hasil penelitian ini, saran yang diberikan kepada PT. Pertamina selaku produsen sekaligus distributor Bahan Bakar Minyak (BBM) adalah meningkatkan promosi dan *creative marketing* kepada masyarakat di wilayah Kabupaten/Kota yang memiliki tingkat konsumsi BBM jenis Premium dan Solar yang masih tinggi, agar dapat beralih menggunakan BBM dengan kualitas yang lebih baik seperti Pertamax dan Pertamina Dex.
2. Pada penelitian selanjutnya diharapkan dapat dilakukan penelitian lebih lanjut tentang pengelompokkan dengan metode *model-based clustering* pada data yang tidak berdistribusi t multivariat dan dapat diterapkan pada data panel.

DAFTAR PUSTAKA

- Agustini, M. (2017). *Model-Based Clustering dengan Distribusi t Multivariat Menggunakan Kriteria Integrated Completed Likelihood dan Minimum Message Length Studi Kasus: Pengelompokan Provinsi di Indonesia Menurut Indikator Pasar Tenaga Kerja Tahun 2012-2015*. Tesis. Departemen Statistika Institut Teknologi Sepuluh Nopember.
- Alamsyah & Anshori, A. (2004). *Analisa Formulasi Konsumsi Bahan Bakar pada Lalu Lintas sebagai Fungsi dari Berhenti Tundaan Waktu Perjalanan Frekuensi Berhenti Frekuensi Tundaan dan Kemacetan Kendaraan Studi Kasus di Kotamadya Malang*. Tesis. Departemen Teknik Sipil Institut Teknologi Sepuluh Nopember.
- Andesberg, M.R. (1973). *Cluster Analysis for Applications*. New York: Academic Press.
- Andrews, J.L., McNicholas, P.D. & Subedi, S. (2011), "Model-based classification via mixtures of multivariate t distributions", *Computational Statistics and Data Analysis*, 55(1), 520-529.
- Arellano, R.B., Valle, & Genton, M.G. (2010). "Multivariate extended skew-t distributions and related families", *International Journal of Statistics*, 68(3), 201-234.
- Banfield, J.D. & Raftery, A.E. (1993), "Model-Based Gaussian and non-Gaussian Clustering", *Biometrics*, 49(3), 803-821.
- Bathke, A., Friedrich, S., Konietzschke, F., Pauly, M., Staffen, W., Strobl, N. & Holler, Y. (2016). "Using EEG, SPECT, and Multivariate Resampling Methods to Differentiate Between Alzheimer's and other Cognitive Impairments". *arXiv*, 8-11, DOI:1606.09004v1.
- Biernacki, C., Celeux, G. & Govaert, G. (2000), "Assessing a Mixture Model for Clustering with the Integrated Completed Likelihood", *IEEE Trans*, 719-725.

- BPH Migas. (2017). *Konsumsi BBM Nasional*. Diakses pada 14 Februari 2017, dari Badan Pengatur Hilir Minyak dan Gas Bumi Indonesia: <http://www.bphmigas.go.id/konsumsi-bbmnasional>.
- BPS. (2016). *Jumlah Kendaraan Bermotor Nasional Tahun 1949-2013*. Diakses pada 14 Februari 2017, dari Badan Pusat Statistik: <https://www.bps.go.id/index.php/linkTabelStatistik/1425>.
- DeCarlo, L.T. (1997). "On the Meaning and Use of Kurtosis". *Psychological Methods*, 2(3), 292-307.
- Dempster, A.P., Laird, N.M. & Rubin, D.B. (1977), "Maximum Likelihood for Incomplete Data via the EM Algorithm (with discussion)". *Journal of the Royal Statistical Society*, 39, 1-38.
- Hair, J.F., Black, W.C., Babin, B.J. & Anderson, R.E. (2009). *Multivariate Data Analysis*. USA: Prentice Hall.
- Handajani, M. (2012), "Model Konsumsi Bahan Bakar Minyak (BBM) Akibat Sistem Transportasi Kota di Jawa", *Prosiding Konferensi Nasional Pascasarjana Teknik Sipil (KNPTS)*, 83.
- Hosseini, S. (2010), "Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty", *Expert Systems with Applications*, 5259-5264.
- Johnson, R.A. & Wichern, D.W. (2007). *Applied Multivariate Statistical Analysis 6th Edition*. New Jersey: Pearson Prentice Hall.
- Kamber. (2006). *Data Mining: Concept and Techniques, Second Edition*. Waltham: Morgan Kaufmann Publishers.
- Kementerian ESDM. (2014). *Realisasi BBM per Provinsi*. Diakses pada 14 Februari 2017, dari Kementerian Energi dan Sumber Daya Mineral Indonesia: <http://kip.esdm.go.id/pusdatin/index.php/datainformasi/data-energi/minyak-dan-gas-bumi/realisasi-bbmperprovinsi>.
- Koesoemadinata, R. P. (1980). *Geologi Minyak Gasbumi*. Bandung: Institut Teknologi Bandung.

- McLachlan, G.J. & Peel, D. (2000). *Finite Mixture Models*. New York: John Wiley and Sons.
- Nuryanti. (2007). *Analisis Karakteristik Konsumsi Energi pada Sektor Rumah Tangga di Indonesia*. Tugas Akhir. Sekolah Tinggi Teknologi Nuklir Batan.
- Pardede, T. (2008), "Perbandingan metode berbasis model (model-based) dengan metode-metode K-Means dalam analisis gugus", *Jurnal Sigma, Sains dan Teknologi*, 11(2), 157-166.
- Pertamina. (2016). *Laporan Realisasi BBM Jawa Timur Tahun 2014-2016*. Surabaya: Retail Fuel Marketing Pertamina UPMS V.
- Rencher, A.C. & Christensen, W.F. (2002). *Methods of Multivariate Analysis*. Canada: John Wiley & Sons, Inc.
- Rousseeuw, P.J, & Van Zomeren, B.C . (1990), "Unmasking multivariate outliers and leverage points", *Journal of the American Statistical Association*, 85, 633-639.
- Susanty, S.E. (2012). *Segmentasi Pelanggan Menggunakan Two Stage Clustering dan LRFM Model pada Divisi Marketing PT.XYZ*. Tugas Akhir. Departemen Sistem Informasi Institut Teknologi Sepuluh Nopember.
- Wahyudi, S. H. (2010). *Analisis Konsumsi dan Perilaku Konsumen dalam Penggunaan Energi Bahan Bakar Minyak untuk Kendaraan Bermotor di Surakarta*. Tugas Akhir. Jurusan Teknik Industri Universitas Sebelas Maret.
- Walpole, R.E. (1995). *Pengantar Metode Statistika*. Alih Bahasa: Ir. Bambang Sumantri, Jakarta : PT Gramedia Pustaka Utama.

(Halaman ini sengaja dikosongkan)

LAMPIRAN

Lampiran 1 Data Jumlah Konsumsi BBM serta Faktor-Faktor yang Mempengaruhi pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2011 (X1 sampai X16)

Kode Wilayah	Wilayah	X1	X2	X3	...	X16
1	Kab. Pacitan	25832	16	0	...	119
2	Kab. Ponorogo	66781	120	0	...	0
3	Kab. Trenggalek	39348	168	0	...	1038
4	Kab. Tulungagung	96317.5	536	72	...	413
5	Blitar	118824	888	112	...	10
6	Kediri	170894.5	1384	0	...	0
7	Kab. Malang & Batu	229686	960	1632	...	396
8	Kab. Lumajang	69334	400	0	...	0
9	Kab. Jember	181336	1176	0	...	713
10	Kab. Banyuwangi	136064	464	0	...	1182
11	Kab. Bondowoso	42976	200	0	...	0
12	Kab. Situbondo	48872	216	0	...	303
13	Probolinggo	94093	528	0	...	532
14	Pasuruan	163284	1184	88	...	103
15	Kab. Sidoarjo	299112	4880	1232	...	0
16	Mojokerto	120136	680	112	...	0
17	Kab. Jombang	107979	728	288	...	0

Lampiran 1 Data Jumlah Konsumsi BBM serta Faktor-Faktor yang Mempengaruhi pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2011 (X1 sampai X16)
(Lanjutan)

Kode Wilayah	Wilayah	X1	X2	X3	...	X16
18	Kab. Nganjuk	76658.5	424	0	...	0
19	Madiun	85516	640	0	...	0
20	Kab. Magetan	59276	248	0	...	0
21	Kab. Ngawi	51464.5	232	16	...	0
22	Kab. Bojonegoro	73482	296	104	...	0
23	Kab. Tuban	69250	312	24	...	1062
24	Kab. Lamongan	85627	440	0	...	5118
25	Kab. Gresik	135938	1320	40	...	1360
26	Kab. Bangkalan	58258	320	88	...	181
27	Kab. Sampang	31128	88	0	...	421
28	Kab. Pamekasan	54370	272	0	...	89
29	Kab. Sumenep	51410	144	0	...	4583
32	Kota Malang	152465.6	2952	728	...	0
37	Kota Surabaya	531486	13282	7448	...	0

Lampiran 2 Data Jumlah Konsumsi BBM serta Faktor-Faktor yang Mempengaruhi pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2012 (X1 sampai X16)

Kode Wilayah	Wilayah	X1	X2	X3	...	X16
1	Kab. Pacitan	28973	61	0	...	122
2	Kab. Ponorogo	74380	172	0	...	0
3	Kab. Trenggalek	45054	158	0	...	1029
4	Kab. Tulungagung	104580.5	728	84	...	393
5	Blitar	131247	937	120	...	11
6	Kediri	189406.5	1257	40	...	0
7	Kab. Malang & Batu	254847	1389	1124	...	398
8	Kab. Lumajang	80049	271	0	...	0
9	Kab. Jember	203616	1263	0	...	1427
10	Kab. Banyuwangi	151832	728	0	...	927
11	Kab. Bondowoso	48584	232	0	...	0
12	Kab. Situbondo	55280	328	0	...	607
13	Probolinggo	105749	606	0	...	562
14	Pasuruan	187319	1484	108	...	95
15	Kab. Sidoarjo	325503	5550	1236	...	0
16	Mojokerto	133030	771	100	...	0
17	Kab. Jombang	119799	726	312	...	0

Lampiran 2 Data Jumlah Konsumsi BBM serta Faktor-Faktor yang Mempengaruhi pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2012 (X1 sampai X16) (Lanjutan)

Kode Wilayah	Wilayah	X1	X2	X3	...	X16
18	Kab. Nganjuk	87957.5	476	0	...	0
19	Madiun	95249	648	0	...	0
20	Kab. Magetan	67271	326	0	...	0
21	Kab. Ngawi	57044.5	221	16	...	0
22	Kab. Bojonegoro	85813	386	100	...	0
23	Kab. Tuban	78021	344	28	...	2125
24	Kab. Lamongan	98199	489	0	...	5119
25	Kab. Gresik	154318	1516	48	...	2721
26	Kab. Bangkalan	65016	376	16	...	181
27	Kab. Sampang	35648	136	0	...	553
28	Kab. Pamekasan	60968	380	0	...	89
29	Kab. Sumenep	57485	188	0	...	5244
32	Kota Malang	158185.4	2696	1152	...	0
37	Kota Surabaya	572529	14220	7832	...	0

Lampiran 3 Data Jumlah Konsumsi BBM serta Faktor-Faktor yang Mempengaruhi pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2013 (X1 sampai X16)

Kode Wilayah	Wilayah	X1	X2	X3	...	X16
1	Kab. Pacitan	30976	80	0	...	124
2	Kab. Ponorogo	77928	380	0	...	0
3	Kab. Trenggalek	46668	248	0	...	1019
4	Kab. Tulungagung	108624	1627	224	...	374
5	Blitar	135736	1807	168	...	11
6	Kediri	194668	2038	64	...	0
7	Kab. Malang & Batu	265472	2273	1240	...	401
8	Kab. Lumajang	83848	416	0	...	0
9	Kab. Jember	211672	2224	0	...	2140
10	Kab. Banyuwangi	157344	1360	0	...	672
11	Kab. Bondowoso	50072	432	0	...	0
12	Kab. Situbondo	57808	656	0	...	910
13	Probolinggo	109488.4	901	0	...	592
14	Pasuruan	199360	2348	132	...	86
15	Kab. Sidoarjo	338776	8530	1448	...	0
16	Mojokerto	140072	1418	128	...	0
17	Kab. Jombang	122320	1174	352	...	0

Lampiran 3 Data Jumlah Konsumsi BBM serta Faktor-Faktor yang Mempengaruhi pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2013 (X1 sampai X16)
(Lanjutan)

Kode Wilayah	Wilayah	X1	X2	X3	...	X16
18	Kab. Nganjuk	90736	935	12	...	0
19	Madiun	97064	861	0	...	0
20	Kab. Magetan	69704	356	0	...	0
21	Kab. Ngawi	61672	284	8	...	0
22	Kab. Bojonegoro	91464	669	216	...	0
23	Kab. Tuban	81144	540	0	...	3187
24	Kab. Lamongan	104064	861	0	...	5119
25	Kab. Gresik	161376	2594	64	...	4081
26	Kab. Bangkalan	66432	452	4	...	181
27	Kab. Sampang	39320	136	0	...	685
28	Kab. Pamekasan	62920	472	0	...	89
29	Kab. Sumenep	57876	208	0	...	5904
32	Kota Malang	160360	4218	1480	...	0
37	Kota Surabaya	577088	21266	8214	...	0

Lampiran 4 Data Jumlah Konsumsi BBM serta Faktor-Faktor yang Mempengaruhi pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2014 (X1 sampai X16)

Kode Wilayah	Wilayah	X1	X2	X3	...	X16
1	Kab. Pacitan	32024	289.5	0	...	108
2	Kab. Ponorogo	80432	858	0	...	0
3	Kab. Trenggalek	48260	512	0	...	1019
4	Kab. Tulungagung	110984	2571	472	...	0
5	Blitar	139232	3237	108	...	12
6	Kediri	196752	3207	196	...	0
7	Kab. Malang & Batu	273816	3487.5	1408	...	410
8	Kab. Lumajang	85125.76	681	0	...	0
9	Kab. Jember	215384	3444	0	...	2140
10	Kab. Banyuwangi	162600	2344	0	...	740
11	Kab. Bondowoso	50912	600	0	...	0
12	Kab. Situbondo	58576	1012	0	...	910
13	Probolinggo	112946.8	1330	0	...	635
14	Pasuruan	207136	3571	156	...	128
15	Kab. Sidoarjo	336520	12636	1528	...	0
16	Mojokerto	143928	2164	156	...	0
17	Kab. Jombang	124016	1686	508	...	0

Lampiran 4 Data Jumlah Konsumsi BBM serta Faktor-Faktor yang Mempengaruhi pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2014 (X1 sampai X16) (Lanjutan)

Kode Wilayah	Wilayah	X1	X2	X3	...	X16
18	Kab. Nganjuk	92680	1351	124	...	0
19	Madiun	98088	1327	0	...	0
20	Kab. Magetan	70016	859	0	...	0
21	Kab. Ngawi	62216	580	0	...	0
22	Kab. Bojonegoro	93776	1112	248	...	0
23	Kab. Tuban	83248	804	0	...	3688
24	Kab. Lamongan	109048	1446	0	...	5119
25	Kab. Gresik	167203	4089	60	...	4136
26	Kab. Bangkalan	69552	732	12	...	980
27	Kab. Sampang	39432	308	0	...	1120
28	Kab. Pamekasan	66296	848	0	...	96
29	Kab. Sumenep	60884	280	0	...	5983
32	Kota Malang	159144	6084	1536	...	0
37	Kota Surabaya	576144	30114	9572	...	0

Lampiran 5 Data Jumlah Konsumsi BBM serta Faktor-Faktor yang Mempengaruhi pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2015 (X1 sampai X16)

Kode Wilayah	Wilayah	X1	X2	X3	...	X16
1	Kab. Pacitan	28768	3398	0	...	92
2	Kab. Ponorogo	73632	9191. 5	0	...	0
3	Kab. Trenggalek	46192	5296	0	...	1019
4	Kab. Tulungagung	99480	16118 .25	992	...	0
5	Blitar	126024	18510	160	...	13
6	Kediri	180496	18152	268	...	0
7	Kab. Malang & Batu	248928	24075 .91	333 2	...	419
8	Kab. Lumajang	77888	9078	0	...	0
9	Kab. Jember	191440	23512	0	...	2140
10	Kab. Banyuwangi	136776	30512	0	...	808
11	Kab. Bondowoso	46368	4112	0	...	0
12	Kab. Situbondo	52752	7016	16	...	910
13	Probolinggo	105800	9150	0	...	678
14	Pasuruan	190872	18940	324	...	170
15	Kab. Sidoarjo	298496	42433	303 2	...	0
16	Mojokerto	132584	12548	392	...	0
17	Kab. Jombang	118384	8004	708	...	0

Lampiran 5 Data Jumlah Konsumsi BBM serta Faktor-Faktor yang Mempengaruhi pada Kabupaten/Kota di Provinsi Jawa Timur Tahun 2015 (X1 sampai X16) (Lanjutan)

Kode Wilayah	Wilayah	X1	X2	X3	...	X16
18	Kab. Nganjuk	86784	7254	120	...	0
19	Madiun	89176	9240	0	...	0
20	Kab. Magetan	63592	7123	0	...	0
21	Kab. Ngawi	57600	5787	0	...	0
22	Kab. Bojonegoro	89792	7374	1312	...	0
23	Kab. Tuban	74576	8185	0	...	3989
24	Kab. Lamongan	102672	10960	0	...	5119
25	Kab. Gresik	152024	19309	124	...	4191
26	Kab. Bangkalan	67296	5504	272	...	1179
27	Kab. Sampang	37824	3424	0	...	1555
28	Kab. Pamekasan	60520	7616	0	...	103
29	Kab. Sumenep	58142	2844	0	...	6062
32	Kota Malang	143653.4	24919	2180	...	0
37	Kota Surabaya	494512	90382	17788	...	0

Lampiran 6 *Output Deteksi Outlier Subset Data s₁2011 menggunakan Software R*

```
> library(mvoutlier)
> rekap2011=read.csv(file.choose(),header=T, sep=";")
> data2011=data.frame(rekap2011[3:7])
> dd.plot(data2011, quan=0.9, alpha=0.05)

$outliers
[1] FALSE FALSE FALSE FALSE FALSE FALSE TRUE FALSE TRUE FALSE FALSE FALSE
[13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[25] FALSE FALSE FALSE FALSE FALSE FALSE TRUE

$md.cla
[1] 1.4362282 1.1771119 1.1340319 1.7919650 1.8332163 3.1728612 3.5188237
[8] 0.8613965 3.7542097 3.5741719 1.9103646 1.3997982 1.1698869 2.3257713
[15] 2.8700768 1.3886613 1.1724510 1.0379492 1.2938826 1.4149607 1.1557135
[22] 2.6149878 1.9778710 1.6303295 1.7063273 1.6713831 1.6883453 1.2234410
[29] 1.3466510 3.1369131 4.8480149

$md.rob
[1] 1.947281 1.065212 1.299349 1.611352 1.789287 3.181380 6.999565
[8] 1.065480 7.687752 2.999168 1.703805 1.249272 2.280036 2.668857
[15] 3.308615 1.451372 1.586329 1.184984 1.144811 1.376205 1.033693
[22] 2.251241 1.819598 1.472116 2.699461 1.445345 1.518589 1.106507
[29] 1.257309 2.858357 10.441701
```

Lampiran 7 Kabupaten/Kota yang terdeteksi sebagai *outlier* pada subset data s_1 tahun 2011-2015

Kode Wilayah	Wilayah	Tahun				
		2011	2012	2013	2014	2015
1	Kab. Pacitan					
2	Kab. Ponorogo					
3	Kab. Trenggalek					
4	Kab. Tulungagung					
5	Blitar					
6	Kediri					
7	Kab. Malang & Batu	(*)	(*)	(*)	(*)	(*)
8	Kab. Lumajang					
9	Kab. Jember	(*)	(*)	(*)	(*)	(*)
10	Kab. Banyuwangi					
11	Kab. Bondowoso					
12	Kab. Situbondo					
13	Probolinggo					
14	Pasuruan					
15	Kab. Sidoarjo					
16	Mojokerto					
17	Kab. Jombang					
18	Kab. Nganjuk					
19	Madiun					
20	Kab. Magetan					
21	Kab. Ngawi					
22	Kab. Bojonegoro					
23	Kab. Tuban					
24	Kab. Lamongan					

Lampiran 7 Kabupaten/Kota yang terdeteksi sebagai *outlier* pada *subset* data s_1 tahun 2011-2015 (Lanjutan)

Kode Wilayah	Wilayah	Tahun				
		2011	2012	2013	2014	2015
25	Kab. Gresik					
26	Kab. Bangkalan					
27	Kab. Sampang					
28	Kab. Pamekasan					
29	Kab. Sumenep					
32	Kota Malang					
37	Kota Surabaya	(*)	(*)	(*)	(*)	(*)

Keterangan :

(*) Kabupaten/Kota yang terdeteksi sebagai *outlier* pada *quantile* = 0.9

Lampiran 8 Kabupaten/Kota yang terdeteksi sebagai *outlier* pada subset data s_2 tahun 2011-2015

Kode Wilayah	Wilayah	Tahun				
		2011	2012	2013	2014	2015
1	Kab. Pacitan					
2	Kab. Ponorogo					
3	Kab. Trenggalek					
4	Kab. Tulungagung					
5	Blitar					
6	Kediri					
7	Kab. Malang & Batu		(*)	(*)	(*)	(*)
8	Kab. Lumajang					
9	Kab. Jember		(*)	(*)	(*)	(*)
10	Kab. Banyuwangi					
11	Kab. Bondowoso					
12	Kab. Situbondo					
13	Probolinggo					
14	Pasuruan					
15	Kab. Sidoarjo					
16	Mojokerto					
17	Kab. Jombang					
18	Kab. Nganjuk					
19	Madiun					
20	Kab. Magetan					
21	Kab. Ngawi					
22	Kab. Bojonegoro					
23	Kab. Tuban					
24	Kab. Lamongan	(*)				

Lampiran 8 Kabupaten/Kota yang terdeteksi sebagai *outlier* pada *subset* data s_2 tahun 2011-2015 (Lanjutan)

Kode Wilayah	Wilayah	Tahun				
		2011	2012	2013	2014	2015
25	Kab. Gresik					
26	Kab. Bangkalan					
27	Kab. Sampang					
28	Kab. Pamekasan					
29	Kab. Sumenep	(*)				
32	Kota Malang					
37	Kota Surabaya	(*)	(*)	(*)	(*)	(*)

Keterangan :

(*) Kabupaten/Kota yang terdeteksi sebagai *outlier* pada *quantile* = 0.9

Lampiran 9 Kabupaten/Kota yang terdeteksi sebagai *outlier* pada subset data s_3 tahun 2011-2015

Kode Wilayah	Wilayah	Tahun				
		2011	2012	2013	2014	2015
1	Kab. Pacitan					
2	Kab. Ponorogo					
3	Kab. Trenggalek					
4	Kab. Tulungagung					
5	Blitar					
6	Kediri					
7	Kab. Malang & Batu	(*)				
8	Kab. Lumajang					
9	Kab. Jember					
10	Kab. Banyuwangi					
11	Kab. Bondowoso					
12	Kab. Situbondo					
13	Probolinggo					
14	Pasuruan					
15	Kab. Sidoarjo	(*)	(*)	(*)	(*)	(*)
16	Mojokerto					
17	Kab. Jombang					
18	Kab. Nganjuk					
19	Madiun					
20	Kab. Magetan					
21	Kab. Ngawi					
22	Kab. Bojonegoro					
23	Kab. Tuban					
24	Kab. Lamongan					

Lampiran 9 Kabupaten/Kota yang terdeteksi sebagai *outlier* pada *subset* data s_3 tahun 2011-2015 (Lanjutan)

Kode Wilayah	Wilayah	Tahun				
		2011	2012	2013	2014	2015
25	Kab. Gresik					
26	Kab. Bangkalan					
27	Kab. Sampang					
28	Kab. Pamekasan					
29	Kab. Sumenep					
32	Kota Malang		(*)	(*)		
37	Kota Surabaya	(*)	(*)	(*)	(*)	(*)

Keterangan :

(*) Kabupaten/Kota yang terdeteksi sebagai *outlier* pada *quantile* = 0.9

Lampiran 10 Kabupaten/Kota yang terdeteksi sebagai *outlier* pada *subset* data s_4 tahun 2011-2015

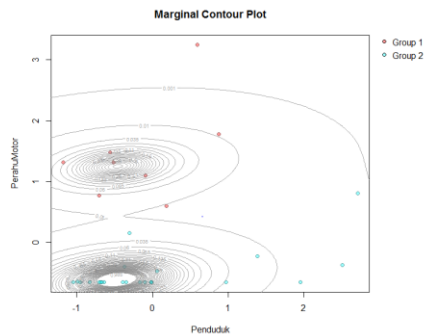
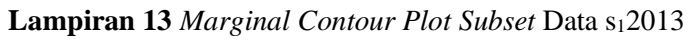
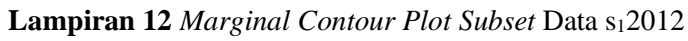
Kode Wilayah	Wilayah	Tahun				
		2011	2012	2013	2014	2015
1	Kab. Pacitan					
2	Kab. Ponorogo					
3	Kab. Trenggalek					
4	Kab. Tulungagung					
5	Blitar					
6	Kediri					(*)
7	Kab. Malang & Batu	(*)				
8	Kab. Lumajang					
9	Kab. Jember					
10	Kab. Banyuwangi					
11	Kab. Bondowoso					
12	Kab. Situbondo					
13	Probolinggo					
14	Pasuruan			(*)		(*)
15	Kab. Sidoarjo		(*)	(*)	(*)	
16	Mojokerto					
17	Kab. Jombang					
18	Kab. Nganjuk					
19	Madiun		(*)			
20	Kab. Magetan					
21	Kab. Ngawi					
22	Kab. Bojonegoro					
23	Kab. Tuban					
24	Kab. Lamongan					

Lampiran 10 Kabupaten/Kota yang terdeteksi sebagai *outlier* pada *subset* data s_4 tahun 2011-2015 (Lanjutan)

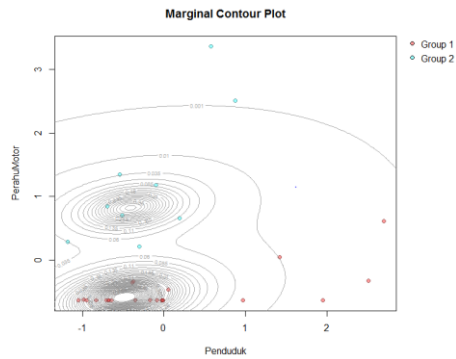
Kode Wilayah	Wilayah	Tahun				
		2011	2012	2013	2014	2015
25	Kab. Gresik	(*)			(*)	
26	Kab. Bangkalan					
27	Kab. Sampang					
28	Kab. Pamekasan					
29	Kab. Sumenep					
32	Kota Malang					
37	Kota Surabaya	(*)	(*)	(*)	(*)	(*)

Keterangan :

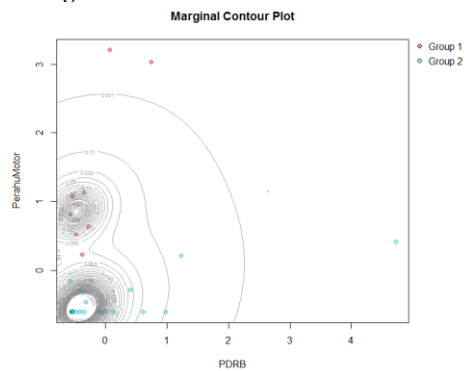
(*) Kabupaten/Kota yang terdeteksi sebagai *outlier* pada *quantile* = 0.9



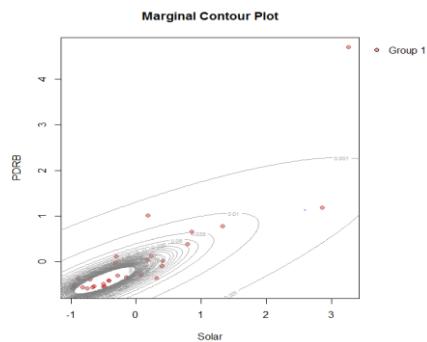
Lampiran 14 *Marginal Contour Plot Subset Data s₁2014*

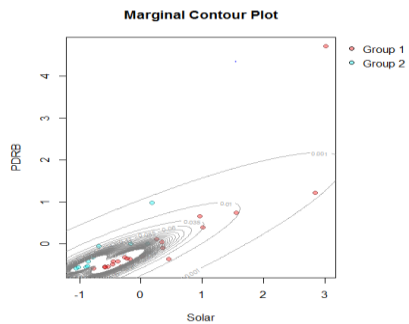
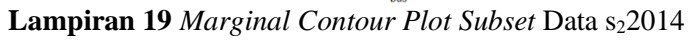
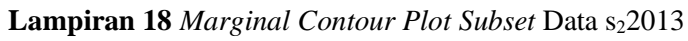


Lampiran 15 *Marginal Contour Plot Subset Data s₁2015*

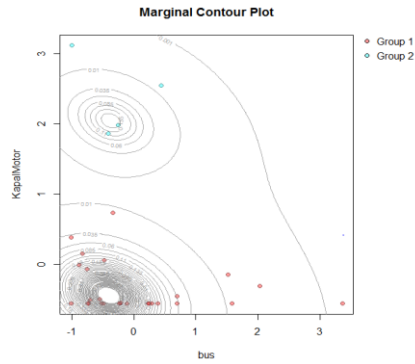


Lampiran 16 *Marginal Contour Plot Subset Data s₂2011*

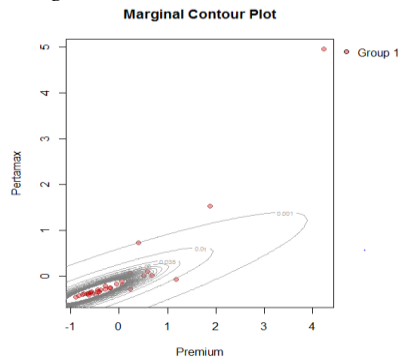




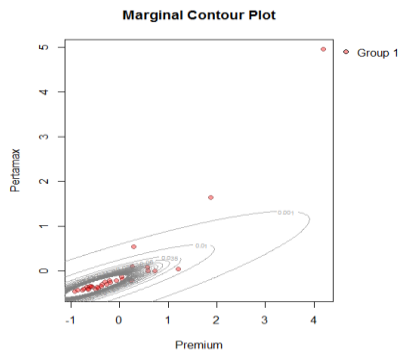
Lampiran 20 *Marginal Contour Plot Subset Data s₂2015*



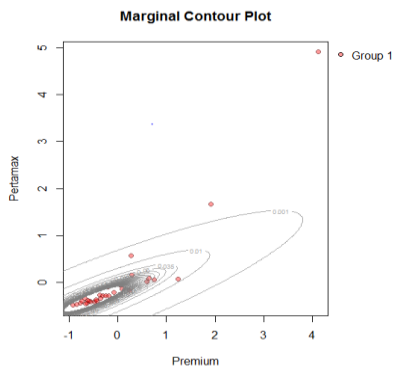
Lampiran 21 *Marginal Contour Plot Subset Data s₃2011*



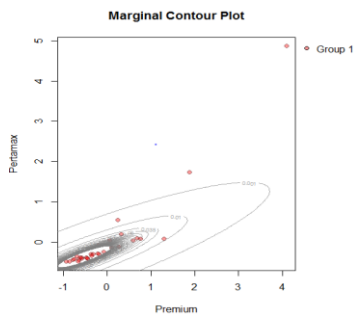
Lampiran 22 *Marginal Contour Plot Subset Data s₃2012*



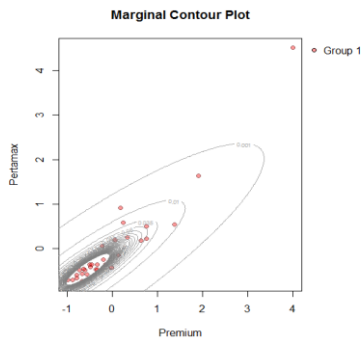
Lampiran 23 *Marginal Contour Plot Subset Data s_3 2013*



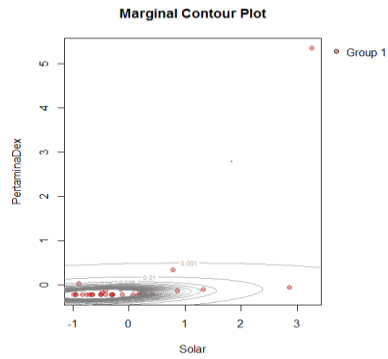
Lampiran 24 *Marginal Contour Plot Subset Data s_3 2014*



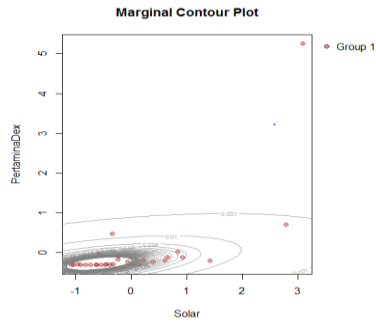
Lampiran 25 *Marginal Contour Plot Subset Data s_3 2015*



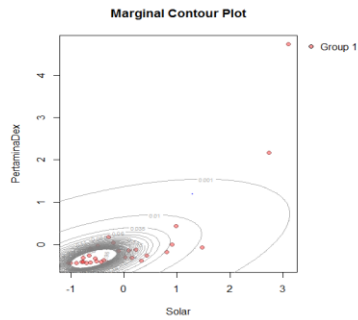
Lampiran 26 *Marginal Contour Plot Subset Data s42011*

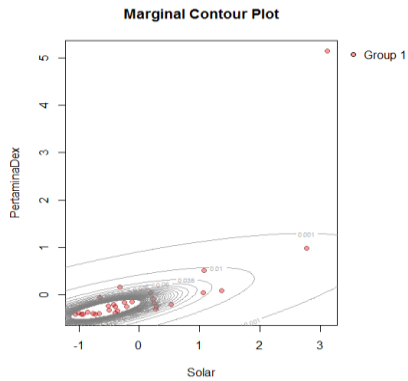
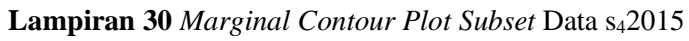


Lampiran 27 *Marginal Contour Plot Subset Data s42012*



Lampiran 28 *Marginal Contour Plot Subset Data s42013*





Lampiran 31 Output Nilai ICL pada Subset Data s₁2015

```

$allicl
      G=1      G=2  G=3  G=4  G=5  G=6  G=7  G=8  G=9
UUUU -245.8583 -274.5478 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UUUC      -Inf -276.7127 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CUCU      -Inf -223.9065 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CUCC      -Inf -221.0396 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CUUU      -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CUUC      -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CCCU      -Inf -239.5695 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CCCC      -Inf -236.5946 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CIUU -333.6299 -326.2140 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CIUC      -Inf -275.7887 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CICU      -Inf -314.5913 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CICC      -Inf -318.4446 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UIIU -332.8351 -320.8634 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UIIC      -Inf -318.8036 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CIIU      -Inf -324.2961 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CIIC      -Inf -319.0157 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UIUU      -Inf -321.5620 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UIUC      -Inf -320.8803 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UCCU      -Inf -235.4375 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UCCC      -Inf -237.6992 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UUCU      -Inf -261.8857 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UUCC      -Inf -258.6190 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UICU      -Inf -310.0497 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UICC      -Inf -309.6193 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UCUU      -Inf -355.9093 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UCUC      -Inf -355.2245 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CCUU      -Inf -313.5321 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CCUC      -Inf -310.1044 -Inf -Inf -Inf -Inf -Inf -Inf -Inf

$bestmodel
[1] "The best model (ICL of -221.04) is CUCC with G=2"

$classification
[1] 1 1 1 1 1 2 2 1 2 1 1 1 1 2 2 1 1 1 1 2 2 1 2 1 1 1
1 2 2

$modelname
[1] "CUCC"

```

Lampiran 32 *Output* Estimasi Parameter *Subset* Data s₁2015 pada
Jumlah Kelompok (G) = 2 dengan Model CUCC

```
$parameters$df
[1] 2.00075 2.00075

$parameters$mean
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] -0.52681203 -0.384665 -0.58493044 -0.4715837 -0.2703281
[2,]  0.07754443  0.153339  0.09164362  0.1155500 -0.5832636

$parameters$lambda
[1] 0.03609328 0.03609328

$parameters$d
, , 1

      [,1]      [,2]      [,3]      [,4]
[1,] -0.08267082 -0.5063487 -0.05597751  0.74246439
0.4270709897
[2,]  0.15465114 -0.6442548  0.67538848 -0.29704089 -
0.1289800386
[3,] -0.26402171 -0.5442534 -0.65668950 -0.45037695
0.0005153305
[4,] -0.07920232 -0.1472248 -0.11879194  0.39652795 -
0.8948217133
[5,] -0.94512861  0.1032460  0.30881109 -0.02096486
0.0163815530

, , 2

      [,1]      [,2]      [,3]      [,4]
[1,]  0.5252525  0.29394086  0.73488579 -0.28889026
0.11913795
[2,]  0.4179658  0.31731571 -0.07716565  0.67498837 -
0.51288549
[3,]  0.6484388 -0.73059231 -0.20132582 -0.07111862
0.01311792
[4,]  0.3436624  0.52477708 -0.64260550 -0.40813318
0.16428825
[5,]  0.1041043  0.06126491 -0.02268097  0.53786854
0.83402153
```

Lampiran 32 *Output* Estimasi Parameter *Subset* Data s₁2015 pada
Jumlah Kelompok (G) = 2 dengan Model CUCC
(Lanjutan)

\$parameters\$a

, , 1

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	15.54247	0.000000	0.0000000	0.0000000	0.00000000
[2,]	0.00000	6.665339	0.0000000	0.0000000	0.00000000
[3,]	0.00000	0.000000	0.8190501	0.0000000	0.00000000
[4,]	0.00000	0.000000	0.0000000	0.2166353	0.00000000
[5,]	0.00000	0.000000	0.0000000	0.0000000	0.05440239

, , 2

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	15.54247	0.000000	0.0000000	0.0000000	0.00000000
[2,]	0.00000	6.665339	0.0000000	0.0000000	0.00000000
[3,]	0.00000	0.000000	0.8190501	0.0000000	0.00000000
[4,]	0.00000	0.000000	0.0000000	0.2166353	0.00000000
[5,]	0.00000	0.000000	0.0000000	0.0000000	0.05440239

\$parameters\$weights

	[,1]	[,2]
[1,]	0.946417142	0.0384978707
[2,]	2.176579750	0.5787244703
[3,]	1.900001219	0.5058189319
[4,]	0.888776599	0.1873192832
[5,]	0.876302410	0.2220153614
[6,]	0.012801221	0.1667212967
[7,]	0.054397881	0.2972028424
[8,]	2.078253206	0.0951930513
[9,]	0.066123413	0.3383705403
[10,]	0.151145953	0.0009312159
[11,]	1.596472976	0.2062734590
[12,]	1.054078776	0.1572735589
[13,]	0.217014194	0.0056376409
[14,]	0.029539279	0.0012903614
[15,]	0.014792131	0.0940112793
[16,]	0.091155890	2.7529175830
[17,]	0.998718633	0.3881770252
[18,]	1.712103016	0.5917315256
[19,]	1.283951991	0.5840335114
[20,]	1.005652542	0.5449429451
[21,]	2.552764503	0.6664781596
[22,]	0.146071801	1.9090353172

Lampiran 32 *Output* Estimasi Parameter *Subset* Data s₁2015 pada
Jumlah Kelompok (G) = 2 dengan Model CUCC
(Lanjutan)

```
[23,] 0.090877508 1.2651016985
[24,] 0.671684103 0.0058912505
[25,] 0.022885763 0.6084149481
[26,] 1.572095406 0.0094311375
[27,] 0.630554110 0.0044245473
[28,] 0.657242749 0.0068748516
```

```
[29,] 0.637422600 0.0141029463
[30,] 0.102472183 0.8202723706
[31,] 0.001070164 0.0433138792
```

```
$parameters$sigma
, , 1
```

```
      [,1]      [,2]      [,3]      [,4]
[ ,5]
[1,] 0.07027558 0.06835704 0.07701481 0.02335543
0.03063599
[2,] 0.06835704 0.12747794 0.04938334 0.01288110 -
0.09178754
[3,] 0.07701481 0.04938334 0.12469971 0.03191616
0.12054406
[4,] 0.02335543 0.01288110 0.03191616 0.01195233
0.03715777
[5,] 0.03063599 -0.09178754 0.12054406 0.03715777
0.50649209
```

```
, , 2
```

```
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] 0.19220017 0.14227366 0.13519254 0.12537124 0.03349462
[2,] 0.14227366 0.12647862 0.09633822 0.11978530 0.03113673
[3,] 0.13519254 0.09633822 0.36552447 0.03683092 0.02695838
[4,] 0.12537124 0.11978530 0.03683092 0.14606858 0.02678799
[5,] 0.03349462 0.03113673 0.02695838 0.02678799 0.01062581
```

```
$parameters$pi
[1] 0.6736453 0.3263547
```

Lampiran 33 Output Nilai ICL pada Subset Data s₂2015

```

$allicl
      G=1      G=2 G=3 G=4 G=5 G=6 G=7 G=8 G=9
UUUU -317.3005  -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UUUC  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CUCU  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CUCC  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CUUU  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CUUC  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CCCU  -Inf -319.4618 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CCCC  -Inf -316.0565 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CIUU -394.2080  -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CIUC  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CICU  -Inf -373.6661 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CICC  -Inf -370.2470 -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UIIU -402.3562  -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UIIC  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CIIU  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CIIC  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UIUU  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UIUC  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UCCU  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UCCC  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UUCU  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UUCC  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UICU  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UICC  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UCUU  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
UCUC  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CCUU  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf
CCUC  -Inf      -Inf -Inf -Inf -Inf -Inf -Inf -Inf -Inf

$bestmodel
[1] "The best model (ICL of -316.06) is CCCC with G=2"

$classification
[1] 2 2 2 2 2 2 2 2 2 2 2 2 1 1 2 2 2 2 2 2 2 2 2 2
2 2 2

$modelname
[1] "CCCC"

```


Lampiran 34 *Output* Estimasi Parameter *Subset* Data s_2 2015 pada
Jumlah Kelompok (G) = 2 dengan Model CCCC

```

$parameters$df
[1] 2 2

$parameters$mean
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] 2.6864447 1.5318811 0.9321492 1.4190416 1.2013773
[2,] -0.5418093
[3,] -0.5146548 -0.4467401 -0.3364533 -0.5360271 -0.4320075
[4,] -0.3500097

$parameters$lambda
[1] Inf

$parameters$d
[1] Inf

$parameters$a
[1] Inf

$parameters$weights
      [,1]      [,2]
[1,] 0.051407091 1.288044631
[2,] 0.049985029 1.979119009
[3,] 0.047991215 1.568358785
[4,] 0.050831588 0.591219559
[5,] 0.043822151 0.664265212
[6,] 0.071304102 0.083521918
[7,] 0.044403401 0.120805129
[8,] 0.057434561 2.784588021
[9,] 0.041705315 0.152969158
[10,] 0.074607851 0.318707096
[11,] 0.048653418 1.361934022
[12,] 0.062996242 1.083531187
[13,] 0.053284333 0.143097037
[14,] 0.226697798 0.053016991
[15,] 3.776247667 0.054660175
[16,] 0.078223507 0.667434816
[17,] 0.056785560 1.954210781
[18,] 0.057355771 2.302206422
[19,] 0.061505457 2.447979347
[20,] 0.051872528 1.628439334
[21,] 0.066870304 1.022038228
[22,] 0.075867770 0.291899320
[23,] 0.062819619 0.256782208
[24,] 0.042637644 0.124450828

```

Lampiran 34 *Output* Estimasi Parameter *Subset* Data s_2 2015 pada
Jumlah Kelompok (G) = 2 dengan Model CCCC
(Lanjutan)

```
[25,] 0.113558950 0.100842613
[26,] 0.045111400 1.189573530
[27,] 0.043686770 1.040207101
[28,] 0.048098900 2.721598836
[29,] 0.031368723 0.130973102
[30,] 0.047420231 0.358099980
[31,] 0.006910881 0.006575812
```

```
$parameters$sigma
, , 1
```

	[,1]	[,2]	[,3]	[,4]
[,5]	[,6]			
[1,]	0.132863060	0.10016268	0.07212989	0.11736883
	0.05660787	0.005625299		
[2,]	0.100162676	0.17003365	0.07822233	0.09974512
	0.04957943	-0.031533970		
[3,]	0.072129892	0.07822233	0.05765863	0.07281648
	0.03994583	-0.019455102		
[4,]	0.117368832	0.09974512	0.07281648	0.19081759
	0.06292397	0.022025552		
[5,]	0.056607867	0.04957943	0.03994583	0.06292397
	0.04059084	0.011405476		
[6,]	0.005625299	-0.03153397	-0.01945510	0.02202555
	0.01140548	0.221673985		

```
, , 2
```

	[,1]	[,2]	[,3]	[,4]
[,5]	[,6]			
[1,]	0.132863060	0.10016268	0.07212989	0.11736883
	0.05660787	0.005625299		
[2,]	0.100162676	0.17003365	0.07822233	0.09974512
	0.04957943	-0.031533970		
[3,]	0.072129892	0.07822233	0.05765863	0.07281648
	0.03994583	-0.019455102		
[4,]	0.117368832	0.09974512	0.07281648	0.19081759
	0.06292397	0.022025552		
[5,]	0.056607867	0.04957943	0.03994583	0.06292397
	0.04059084	0.011405476		
[6,]	0.005625299	-0.03153397	-0.01945510	0.02202555
	0.01140548	0.221673985		

```
$parameters$pig
[1] 0.0708632 0.9291368
```

Lampiran 35 Output Nilai ICL pada Subset Data s_3 2015

\$allicl

[illegible]

```
$bestmodel
```

```
[1] "The best model (ICL of -87.47) is UUUU with G=1"
```

```
$classification
```

[illegible]

\$modelname

```
[1] "UUUUU"
```

Lampiran 36 *Output* Estimasi Parameter *Subset* Data s_3 2015 pada
Jumlah Kelompok (G) = 1 dengan Model UUUU

```
$parameters$df
[1] 2

$parameters$mean
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] -0.5370702 -0.4658009 -0.2749467 -0.4155348 -0.3515582
-0.3551157
      [,7]      [,8]      [,9]     [,10]
[1,] -0.3536913 -0.5826798 -0.4322092 -0.2207131

$parameters$lambda
[1] Inf

$parameters$d
[1] Inf

$parameters$a
[1] Inf

$parameters$weights
      [,1]
[1,] 2.346742155
[2,] 1.073142910
[3,] 3.585232404
[4,] 0.481125973
[5,] 0.701901254
[6,] 0.150920169
[7,] 0.111833769
[8,] 1.774053408
[9,] 0.133930114
[10,] 0.090522359
[11,] 2.408725024
[12,] 1.066610051
[13,] 0.287094784
[14,] 0.191607876
[15,] 0.043618291
[16,] 0.787524398
[17,] 0.772493680
[18,] 2.072411342
[19,] 0.969271676
[20,] 2.175185420
[21,] 2.778452829
[22,] 0.392964105
[23,] 0.757365434
```

Lampiran 36 *Output* Estimasi Parameter *Subset* Data s_3 2015 pada
Jumlah Kelompok (G) = 1 dengan Model UUUU
(Lanjutan)

```
[24,] 0.874362705
[25,] 0.259524803
[26,] 2.116234289
[27,] 0.935664380
[28,] 1.958396262
[29,] 1.235259081
[30,] 0.054868208
[31,] 0.002487041

$parameters$sigma
, , 1
      [,1]      [,2]      [,3]      [,4]
[1,] 0.13035598 0.08315112 0.023442572 0.11745033
0.04834020 0.04940177
[2,] 0.08315112 0.06863890 0.014763914 0.07784518
0.03573559 0.03656014
[3,] 0.02344257 0.01476391 0.013074848 0.02246602
0.01128136 0.01272081
[4,] 0.11745033 0.07784518 0.022466016 0.14970255
0.05409822 0.05013129
[5,] 0.04834020 0.03573559 0.011281359 0.05409822
0.03092250 0.02585290
[6,] 0.04940177 0.03656014 0.012720805 0.05013129
0.02585290 0.02564765
[7,] 0.05180660 0.03722846 0.011839564 0.05336915
0.02767118 0.02514464
[8,] 0.15090896 0.08718596 0.027619990 0.11866654
0.04093165 0.05105185
[9,] 0.07305475 0.04620216 0.013883223 0.06220757
0.02629806 0.02875460
[10,] 0.02042096 0.01662849 -0.001077652 -0.06298693 -
0.03462434 -0.01813073
      [,7]      [,8]      [,9]      [,10]
[1,] 0.05180660 0.15090896 0.07305475 0.020420961
[2,] 0.03722846 0.08718596 0.04620216 0.016628486
[3,] 0.01183956 0.02761999 0.01388322 -0.001077652
[4,] 0.05336915 0.11866654 0.06220757 -0.062986935
[5,] 0.02767118 0.04093165 0.02629806 -0.034624337
[6,] 0.02514464 0.05105185 0.02875460 -0.018130731
[7,] 0.02675185 0.04954648 0.02876639 -0.018997912
[8,] 0.04954648 0.22759261 0.08893151 0.080306315
[9,] 0.02876639 0.08893151 0.06039968 0.019152258
[10,] -0.01899791 0.08030631 0.01915226 0.423371759
```


Lampiran 38 *Output* Estimasi Parameter *Subset* Data s_4 2015 pada
Jumlah Kelompok (G) = 1 dengan Model UUUU

```

$parameters$df
[1] 2

$parameters$mean
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] -0.5518758 -0.312707 -0.4754972 -0.3598249 -0.4472034 -
0.5643933
      [,7]      [,8]
[1,] -0.4519182 -0.3185824

$parameters$lambda
[1] Inf

$parameters$d
[1] Inf

$parameters$a
[1] Inf

$parameters$weights
      [,1]
[1,] 1.528360365
[2,] 2.610454248
[3,] 2.166248457
[4,] 0.827359283
[5,] 0.811110582
[6,] 0.046343208
[7,] 0.155704616
[8,] 3.048897564
[9,] 0.177164959
[10,] 0.254188161
[11,] 1.876504899
[12,] 1.036373877
[13,] 0.166370442
[14,] 0.044849208
[15,] 0.071171688
[16,] 0.174088579
[17,] 1.415895563
[18,] 2.258625415
[19,] 2.654205023
[20,] 1.812546549
[21,] 1.318848493
[22,] 0.241981984
[23,] 0.296370366
[24,] 0.179940945

```

Lampiran 38 *Output* Estimasi Parameter *Subset* Data s_4 2015 pada
Jumlah Kelompok (G) = 1 dengan Model UUUU
(Lanjutan)

```
[25,] 0.057047240
[26,] 1.747738241
[27,] 1.640279881
[28,] 3.352462870
[29,] 0.216228040
[30,] 0.307022797
[31,] 0.003194368
```

```
$parameters$sigma
, , 1
```

	[,1]	[,2]	[,3]	[,4]
[,5]	[,6]			
[1,]	0.1531264371	0.041065507	0.12055496	0.07216941
	0.07809310	0.13246187		
[2,]	0.0410655065	0.020048758	0.03955131	0.02341445
	0.02496719	0.03640989		
[3,]	0.1205549553	0.039551310	0.19936266	0.08262443
	0.06852145	0.11226708		
[4,]	0.0721694056	0.023414450	0.08262443	0.05170589
	0.03841212	0.07475716		
[5,]	0.0780931005	0.024967190	0.06852145	0.03841212
	0.10221122	0.05576360		
[6,]	0.1324618691	0.036409887	0.11226708	0.07475716
	0.05576360	0.21289704		
[7,]	0.0602258193	0.022484042	0.05114740	0.03441749
	0.04046154	0.06708367		
[8,]	0.0001274449	-0.005121385	-0.03653716	-0.02442932
	0.01155585	0.03342138		
	[,7]	[,8]		
[1,]	0.06022582	0.0001274449		
[2,]	0.02248404	-0.0051213854		
[3,]	0.05114740	-0.0365371561		
[4,]	0.03441749	-0.0244293243		
[5,]	0.04046154	0.0115558480		
[6,]	0.06708367	0.0334213795		
[7,]	0.03741700	0.0128925798		
[8,]	0.01289258	0.2878088042		

```
$parameters$pi
[1] 1
```


Lampiran 39 Hasil Pengelompokkan Kabupaten/Kota di Jawa Timur Berdasarkan Karakteristik pada *Subset* Data s_1 tahun 2011-2015

Kode Wilayah	Wilayah	Hasil Pengelompokkan				
		2011	2012	2013	2014	2015
1	Kab. Pacitan	1	1	1	1	1
2	Kab. Ponorogo	1	1	1	1	1
3	Kab. Trenggalek	1	1	1	1	1
4	Kab. Tulungagung	1	1	1	1	1
5	Blitar	1	1	1	1	1
6	Kediri	1	1	1	2	2
7	Kab. Malang & Batu	2	2	2	2	2
8	Kab. Lumajang	1	1	1	1	1
9	Kab. Jember	2	2	2	2	2
10	Kab. Banyuwangi	1	1	1	2	1
11	Kab. Bondowoso	1	1	1	1	1
12	Kab. Situbondo	1	1	1	1	1
13	Probolinggo	1	1	1	1	1
14	Pasuruan	1	1	1	2	1
15	Kab. Sidoarjo	1	1	1	2	2
16	Mojokerto	1	1	1	2	2
17	Kab. Jombang	1	1	1	1	1
18	Kab. Nganjuk	1	1	1	1	1
19	Madiun	1	1	1	1	1
20	Kab. Magetan	1	1	1	1	1
21	Kab. Ngawi	1	1	1	1	1
22	Kab. Bojonegoro	1	1	1	1	2
23	Kab. Tuban	1	1	1	1	2
24	Kab. Lamongan	1	1	1	1	1

Lampiran 39 Hasil Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Karakteristik pada *Subset* Data s_1 tahun 2011-2015 (Lanjutan)

Kode Wilayah	Wilayah	Hasil Pengelompokan				
		2011	2012	2013	2014	2015
25	Kab. Gresik	1	1	1	2	2
26	Kab. Bangkalan	1	1	1	1	1
27	Kab. Sampang	1	1	1	1	1
28	Kab. Pamekasan	1	1	1	1	1
29	Kab. Sumenep	1	1	1	1	1
32	Kota Malang	1	1	1	2	2
37	Kota Surabaya	1	1	1	2	2

Lampiran 40 Hasil Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Karakteristik pada *Subset* Data s_2 tahun 2012-2015

Kode Wilayah	Wilayah	Hasil Pengelompokan			
		2012	2013	2014	2015
1	Kab. Pacitan	2	2	2	2
2	Kab. Ponorogo	2	2	2	2
3	Kab. Trenggalek	2	2	2	2
4	Kab. Tulungagung	2	2	2	2
5	Blitar	2	2	2	2
6	Kediri	2	2	2	2
7	Kab. Malang & Batu	2	2	2	2
8	Kab. Lumajang	2	2	2	2
9	Kab. Jember	2	2	2	2
10	Kab. Banyuwangi	2	2	2	2
11	Kab. Bondowoso	2	2	2	2
12	Kab. Situbondo	2	2	2	2
13	Probolinggo	2	2	2	2
14	Pasuruan	1	1	1	1
15	Kab. Sidoarjo	1	1	1	1
16	Mojokerto	2	2	2	2
17	Kab. Jombang	2	2	2	2
18	Kab. Nganjuk	2	2	2	2
19	Madiun	2	2	2	2
20	Kab. Magetan	2	2	2	2
21	Kab. Ngawi	2	2	2	2
22	Kab. Bojonegoro	2	2	2	2
23	Kab. Tuban	2	2	2	2
24	Kab. Lamongan	2	2	2	2

Lampiran 40 Hasil Pengelompokkan Kabupaten/Kota di Jawa Timur Berdasarkan Karakteristik pada *Subset* Data s₂ tahun 2012-2015 (Lanjutan)

Kode Wilayah	Wilayah	Hasil Pengelompokkan			
		2012	2013	2014	2015
25	Kab. Gresik	2	2	2	2
26	Kab. Bangkalan	2	2	2	2
27	Kab. Sampang	2	2	2	2
28	Kab. Pamekasan	2	2	2	2
29	Kab. Sumenep	2	2	2	2
32	Kota Malang	2	2	2	2
37	Kota Surabaya	2	2	2	2

Lampiran 41 Surat Pernyataan Data

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini, mahasiswa jurusan Statistika FMIPA ITS:

Nama : Muhammad Azhar

NRP : 1313100083

menyatakan bahwa data yang saya gunakan dalam Tugas Akhir/Thesis ini merupakan data sekunder yang diambil dari penelitian/buku/Tugas Akhir/Thesis/publikasi lainnya yaitu:

Sumber : Buku Laporan Tahunan Statistik Provinsi Jawa Timur Tahun 2011-2015, BPS Provinsi Jawa Timur

Keterangan : Data jumlah penduduk, PDRB, perahu motor, dan kapal motor pada Kabupaten/Kota di Jawa Timur Tahun 2011-2015

Surat pernyataan ini dibuat dengan sebenar-benarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui,
Pembimbing Tugas Akhir

Surabaya, 15 Juni 2017



Raden Mohamad Atok, Ph.D
NIP. 19710915 199702 1 001



Muhammad Azhar
NRP. 1313100083

BIODATA PENULIS



Muhammad Azhar adalah anak pertama dari dua bersaudara. Penulis lahir di Sidoarjo pada tanggal 26 Februari 1995. Penulis bertempat tinggal di Desa Rejeni, Kecamatan Krembung, Kabupaten Sidoarjo, Jawa Timur. Pendidikan formal yang telah ditempuh dimulai dari SDN Rejeni (2001-2007), SMPN 1 Krembung (2007-2010), SMAN 1 Krembung (2010-2013). Kemudian penulis menempuh pendidikan Sarjana di

Departemen Statistika FMIPA ITS (2013-2017) melalui jalur SBMPTN. Selama perkuliahan penulis aktif mengikuti kegiatan kepanitiaan diantaranya *Young Engineers and Scientists Summit* ITS, Pekan Raya Statistika ITS, Kesma Expo BEM ITS dll. Selain itu selama kuliah, penulis berkesempatan untuk mengikuti beberapa organisasi antara lain BEM ITS (2014/2015), HIMASTA-ITS (2015/2016), dan IKARSIDS (2013/2014). Saat kuliah, penulis pernah melakukan Kerja Praktek di PT. Pertamina (Persero) MOR V Surabaya. Dengan motto "*if you think that something's right, just do it*". Apabila pembaca ingin berdiskusi, ataupun ingin memberikan kritik maupun saran, dapat melalui *e-mail*: muh.azhar.riy@gmail.com.